



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### **Inferring the star formation histories of massive quiescent galaxies with BAGPIPES: evidence for multiple quenching mechanisms**

**Citation for published version:**

Carnall, AC, McLure, RJ, Dunlop, JS & Davé, R 2018, 'Inferring the star formation histories of massive quiescent galaxies with BAGPIPES: evidence for multiple quenching mechanisms', *Monthly Notices of the Royal Astronomical Society*, vol. 480, no. 4, pp. 4379-4401. <https://doi.org/10.1093/mnras/sty2169>

**Digital Object Identifier (DOI):**

[10.1093/mnras/sty2169](https://doi.org/10.1093/mnras/sty2169)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Monthly Notices of the Royal Astronomical Society

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Inferring the star-formation histories of massive quiescent galaxies with BAGPIPES: Evidence for multiple quenching mechanisms

A. C. Carnall,<sup>1</sup><sup>★</sup> R. J. McLure,<sup>1</sup> J. S. Dunlop,<sup>1</sup> R. Davé,<sup>1</sup>

<sup>1</sup>*SUPA<sup>†</sup>, Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, UK*

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

We present Bayesian Analysis of Galaxies for Physical Inference and Parameter ESTimation, or BAGPIPES, a new PYTHON tool which can be used to rapidly generate complex model galaxy spectra and to fit these to arbitrary combinations of spectroscopic and photometric data using the MULTINEST nested sampling algorithm. We extensively test our ability to recover realistic star-formation histories (SFHs) by fitting mock observations of quiescent galaxies from the MUFASA simulation. We then perform a detailed analysis of the SFHs of a sample of 9289 quiescent galaxies from UltraVISTA with stellar masses,  $M_* > 10^{10} M_\odot$  and redshifts  $0.25 < z < 3.75$ . The majority of our sample exhibit SFHs which rise gradually then quench relatively rapidly, over 1–2 Gyr. This behaviour is consistent with recent cosmological hydrodynamic simulations, where AGN-driven feedback in the low-accretion (jet) mode is the dominant quenching mechanism. At  $z > 1$  we also find a class of objects with SFHs which rise and fall very rapidly, with quenching timescales of  $< 1$  Gyr, consistent with quasar-mode AGN feedback. Finally, at  $z < 1$  we find a population with SFHs which quench more slowly than they rise, over  $> 3$  Gyr, which we speculate to be the result of diminishing overall cosmic gas supply. We confirm the mass-accelerated evolution (downsizing) trend, and a trend towards more rapid quenching at higher stellar masses. However, our results suggest that the latter is a natural consequence of mass-accelerated evolution, rather than a change in quenching physics with stellar mass. We find  $61 \pm 8$  per cent of  $z > 1.5$  massive quenched galaxies undergo significant further evolution by  $z = 0.5$ . BAGPIPES is available at [bagpipes.readthedocs.io](http://bagpipes.readthedocs.io).

**Key words:** galaxies: evolution – galaxies: star formation – methods: statistical

## 1 INTRODUCTION

Recent years have seen rapid advances in the quantity and quality of data available for the study of galaxy evolution. Where once studies of large samples of objects relied solely on photometric data, the advent of multi-object and integral-field spectrographs, and the advancing near-infrared capability of both spectroscopic and photometric instruments mean that increasingly detailed studies can now be performed on statistical samples of objects. This presents new challenges, as ever more complex models must be developed to accurately reproduce observed properties, and more advanced statistical techniques are necessary to deal with larger sample sizes and the higher-dimensional, often highly degenerate parameter spaces of more complex models. Techniques from the field of data science for dealing with large datasets, such as machine learning, are also becoming increasingly valuable to astronomers.

One of the most striking observations of recent years is the bimodal distribution in rest-frame colours which galaxies have been

observed to display across at least the 10 – 11 billion years since redshift,  $z = 2 - 3$  (e.g. [Strateva et al. 2001](#); [Bell et al. 2004](#); [Faber et al. 2007](#); [Williams et al. 2009](#); [Brammer et al. 2009](#); [Whitaker et al. 2011, 2013](#); [Straatman et al. 2014, 2016](#)). When plotted on a colour-magnitude diagram, this bimodality manifests as a diffuse blue cloud and tight red sequence, with a sparsely populated green valley in between. In addition, the same bimodality has been observed in several other properties, such as environment (e.g. [Baldry et al. 2006](#); [Hartley et al. 2010](#); [Chuter et al. 2011](#); [Law-Smith & Eisenstein 2017](#)) and morphology (e.g. [Papovich et al. 2012](#); [Bell et al. 2012](#); [Strazzullo et al. 2013](#); [van der Wel et al. 2014](#)), with blue, star-forming galaxies exhibiting a more disk-like morphology and being less strongly clustered than red, quiescent galaxies, which typically exhibit early-type morphologies.

The observation of this bimodality has led to the search for a quenching mechanism, or mechanisms, capable of shutting down star-formation activity in galaxies. Many potential quenching mechanisms have been proposed, falling broadly into two categories (e.g. [Peng et al. 2015](#)). We refer to these as ejective, associated with the expulsion of the gas reservoir which fuels star formation, and preventative, associated with the supply of new gas to the galaxy being

<sup>★</sup> E-mail: [adamc@roe.ac.uk](mailto:adamc@roe.ac.uk)

<sup>†</sup> Scottish Universities Physics Alliance

shut down. The former results in a rapid halting of star formation, whereas the latter leads to a more gradual decline, or strangulation, as the existing reservoir is depleted.

The most commonly discussed process related to the quenching of massive galaxies is active galactic nucleus (AGN) feedback, in which the accretion of gas onto a galaxy fuels an AGN. This, in turn, either ejects the gas reservoir from the galaxy, or prevents further gas from being accreted, depending on the strength of the feedback. Strong AGN-driven outflows, sometimes called quasar-mode feedback, have been observed at high redshifts (e.g. [Maiolino et al. 2012](#); [Förster Schreiber et al. 2014](#)), and are thought, in some cases, to be triggered by major-merger events. However, it has been shown that galaxies quenched in this way are likely to re-ignite star formation (e.g. [Gabor et al. 2010](#)), rather than remaining quiescent.

Sustained quiescence is thought to require a continuing input of energy to heat the circumgalactic medium, preventing new gas from falling onto the galaxy (e.g. [Gabor et al. 2011](#)). Low-accretion-state AGN feedback, sometimes called radio-mode or jet-mode, has been proposed ([Croton et al. 2006](#)) as the source of this energy input. This mechanism has been shown to significantly improve agreement with the observed properties of the red sequence when implemented in modern hydrodynamic simulations, such as MUFASA ([Davé et al. 2017](#)) and IllustrisTNG ([Nelson et al. 2017](#)).

Another example of an ejective quenching process is the tidal/ram-pressure stripping experienced by galaxies falling into clusters. This process is also capable of causing very rapid quenching of star formation, and is thought to give rise to the environmental dependence of the galaxy colour bimodality (e.g. [Peng et al. 2012](#)). Rapid quenching of satellites is also observed in simulations such as Illustris (e.g. [Diemer et al. 2017](#)).

However, recently [Abramson et al. \(2016\)](#) have shown that no qualitatively different processes are required to have acted on different galaxies to explain the galaxy colour bimodality. In their model, quiescence is the end-point of a single evolutionary track along which all galaxies progress at different rates. Star formation in individual galaxies dies down naturally in the same way as the cosmic star formation rate (SFR) density. Under this paradigm, we should not necessarily expect to observe the imprints of quenching processes in the properties of quiescent galaxies.

In order to understand whether, and to what extent, different potential quenching processes contribute to the termination of star formation, much tighter observational constraints on the way quenching behaves will be necessary. Two fundamental questions which we will seek to address in this work are:

- (i) When did quenched galaxies form their stellar mass?
- (ii) How long did the process of quenching take?

It is already known that a complete answer to these questions must include a description of the dependencies on physical properties such as size and mass, and the effects of different environments. Differences in the properties of quenched galaxies observed at different redshifts may also provide useful insights. In this work we will focus on observed-redshift and stellar-mass dependencies for massive quenched galaxies ( $M_* > 10^{10} M_\odot$ ).

Question (i) has been studied by extracting the star formation histories (SFHs) of local galaxies from the fossil records imprinted in their spectra (e.g. [Heavens et al. 2004](#); [Panter et al. 2007](#); [Thomas et al. 2010](#); [Carson & Nichol 2010](#); [Citro et al. 2016](#)). More recently this analysis has been extended to higher redshifts (e.g. [Moresco et al. 2010](#); [Onodera et al. 2012, 2015](#); [Jørgensen & Chiboucas 2013](#); [Gallazzi et al. 2014](#); [Choi et al. 2014](#); [Lonoce et al. 2014](#); [Fumagalli et al. 2016](#); [Pacifi et al. 2016](#); [Siudek et al. 2017](#)) to

connect local galaxies with their precursors at earlier epochs, and build up a consistent picture of quenching across cosmic time.

Whilst individual results for the derived ages of quenched galaxies are difficult to compare due to differences in the definition of age, a coherent picture has been established, often referred to as downsizing or mass-accelerated evolution, with more massive galaxies typically forming their stellar masses at earlier epochs than their less massive counterparts. An average trend towards later formation for samples observed at lower redshifts is also seen, indicating that the assembly of the red sequence is still ongoing, in agreement with the observed evolution of the galaxy stellar mass function (e.g. [Tomczak et al. 2014](#)).

Question (ii) however is less well understood. [Barro et al. \(2013\)](#) and [Schawinski et al. \(2014\)](#) present evidence for a scheme in which there are both fast and slow tracks towards quiescence, with different mechanisms driving quenching. Massive quiescent galaxies at earlier epochs ( $z \gtrsim 1.5$ ) which are denser systems, colloquially known as ‘red nuggets’, quench rapidly through strong AGN-driven outflows, whereas a second population of less-dense star-forming galaxies quench more gradually as a result of strangulation processes. The denser, earlier-quenching systems are theorised to gradually expand through processes such as minor mergers to leave one population of relatively large quiescent galaxies by the present epoch (e.g. [McLure et al. 2013](#)).

The results of [Abramson et al. \(2016\)](#) however show that it is not necessary to assume a change in quenching mechanism to obtain early-time rapid and late-time slower quenching. This property is reproduced by the model of [Gladders et al. \(2013\)](#), which assumes no physics, as part of a more general relationship, in which galaxies cross the green valley between star-forming and quiescent states in  $\approx 20$  per cent of the age of the Universe at the epoch of their quenching. A similar result is also obtained by [Pacifi et al. \(2016\)](#), although they also argue that their results are inconsistent with a single quenching mechanism acting in all circumstances.

Finally, [Peng et al. \(2015\)](#) argue that a comparison of metallicities between local quiescent and star-forming galaxies strongly favours a slow strangulation of star formation over a timescale of  $\approx 4$  Gyr. Clearly, a direct method which could be demonstrated to reliably constrain in detail the SFHs of individual quiescent galaxies would greatly assist in reconciling these disparate results.

Methods employed to obtain this information rely on modelling the light emitted by galaxies as a function of wavelength in terms of the physical parameters of the system (e.g. [Cid Fernandes et al. 2005](#); [Thomas et al. 2017](#); [Wilkinson et al. 2017](#)). These models are then fitted to observational data, which may consist of spectroscopy and/or photometry (sometimes referred to as the spectral energy distribution; SED). Once the posterior distribution for model parameters has been obtained, nuisance parameters may be marginalised out to obtain constraints on the parameters of interest, such as the SFH. A common approach has been to fit spectroscopic data indirectly, either by data compression (e.g. [Heavens et al. 2000](#)), or by using pre-calibrated spectral indices (e.g. [Worthey et al. 1994](#)).

Recently, the development of new Bayesian statistical techniques, such as advanced Markov chain Monte Carlo (MCMC; e.g. [Goodman & Weare 2010](#); [Foreman-Mackey et al. 2013](#)) and nested sampling algorithms ([Skilling 2006](#), [Feroz & Hobson 2008](#); [Feroz et al. 2009, 2013](#)) has begun to enable the efficient exploration of higher-dimensional parameter spaces to obtain posterior distributions for the parameters of complex models.

A new generation of modern spectral modelling and fitting tools has been developed in order to exploit this e.g. BEAGLE ([Chevallard & Charlot 2016](#)) and PROSPECTOR ([Leja et al. 2017](#); [Johnson](#)

**Table 1.** Mean  $5\sigma$  limiting magnitudes within  $2''$  apertures for the  $1\text{ deg}^2$  twelve band catalogue of Mortlock et al. (2017).

Region	Percentage of area	$u^*$	$g'$	$r'$	$i'$	$z'$	$z'_{\text{Subaru}}$	$Y$	$J$	$H$	$K_s$	IRAC1	IRAC2
Deep strips	40%	27.0	27.1	26.6	26.3	25.4	26.4	25.1	24.9	24.6	24.8	25.3	25.1
Wide strips	60%	27.0	27.1	26.6	26.3	25.4	26.4	24.7	24.4	24.1	23.9	25.3	25.1

et al. in prep). These codes allow on-the-fly generation and fitting of complex, self-consistent models to describe galaxies across continuous parameter spaces. Their models include emission and absorption processes due to the stellar population, ionized gas in H II regions, diffuse dust in the interstellar medium (ISM) and neutral gas in the intergalactic medium (IGM) along our line-of-sight. The ability to explore higher-dimensional parameter spaces also opens up the ability to fit more complex SFHs. This is important because it has been shown that the rigid, exponentially declining SFHs typically employed in SED fitting techniques can introduce significant biases into SFH parameter estimates (see Section 4).

The combination of new statistical techniques with the ever increasing volume and quality of data available means that it is now possible to study the behaviour of quenching processes in unprecedented detail. In this work we analyse photometric data for a large sample of quenched galaxies from the UltraVISTA Survey (McCracken et al. 2012) using Bayesian Analysis of Galaxies for Physical Inference and Parameter ESTimation, or BAGPIPES, a new public galaxy spectral modelling framework and fitting tool written in the PYTHON programming language. BAGPIPES provides a highly intuitive application programming interface (API) for rapid, on-the-fly generation (up to hundreds of models per second) of complex, physically realistic model galaxy spectra across continuous parameter spaces. It also provides a tool, built around the MULTINEST nested sampling algorithm (Feroz & Hobson 2008; Feroz et al. 2009, 2013), which allows for the direct fitting of these models to arbitrary combinations of spectroscopic and photometric data.

We analyse our recovered SFHs for these objects in the context of Questions (i) and (ii) posed above in order to understand the quenching properties of our sample as a function of stellar mass and observed redshift across the majority of cosmic time. In a follow-up paper we will extend our analysis using BAGPIPES to the direct fitting of spectroscopic data from VANDELS (McLure et al. 2018; Pentericci et al. 2018) to obtain better constraints.

The structure of this paper is as follows. In Section 2 we introduce the UltraVISTA dataset we will use to explore quenching. In Section 3 we describe the model generation and fitting methodologies employed by the BAGPIPES code. In Section 4 we consider different models within BAGPIPES, in particular different SFH parameterisations, and test their ability to recover realistic SFHs by fitting mock observations of simulated quenched galaxies from the MUFASA suite of cosmological hydrodynamic simulations (Davé et al. 2016). By this process we define a model which is capable of recovering unbiased estimates of the SFH properties of large samples of quenched galaxies from photometric data, making use of a double-power-law SFH parameterisation. In Section 5, we apply this BAGPIPES model to select a sample of 9289 quenched galaxies with  $M_* > 10^{10} M_\odot$  and redshifts in the range  $0.25 < z < 3.75$  from UltraVISTA. In Section 6 we analyse the properties of their SFHs to constrain the epoch and duration of their quenching as a function of stellar mass and observed redshift. We present our conclusions in Section 7.

All magnitudes are quoted in the AB system and for all cosmological calculations we adopt  $\Omega_M = 0.3$ ,  $\Omega_\Lambda = 0.7$  and  $H_0 = 70\text{ km s}^{-1}\text{ Mpc}^{-1}$ . We distinguish between times,  $t$ , which are meas-

ured forwards from the beginning of the Universe (i.e.  $t(z)$  is the age of the Universe at redshift  $z$ ), and ages,  $a$ , which are measured backwards in time from the redshift of observation,  $t(z_{\text{obs}})$ .

## 2 THE ULTRAVISTA DATA

The UltraVISTA Survey (McCracken et al. 2012) is an ultra-deep imaging program over a  $1.5\text{ deg}^2$  contiguous area of the Cosmological Evolution Survey (COSMOS) field in the near-infrared  $YJHK_s$  bands on the Visible and Infrared Survey Telescope for Astronomy (VISTA). The area is divided into deeper and shallower stripes (UltraVISTA deep and UltraVISTA wide, respectively) which each account for half of the total area. The data utilised here comes from the UltraVISTA DR3 release.

In this work we make use of the  $K$ -band selected catalogue compiled by Mortlock et al. (2017), which covers the  $1\text{ deg}^2$  overlap region between UltraVISTA and the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS; Hudelot et al. 2012) T0007 release, which provides optical  $u^*g'r'i'z'$  band imaging. These datasets are also combined with deep  $z'$ -band imaging from the Subaru telescope (Furusawa et al. 2016) and  $3.6\text{ }\mu\text{m} + 4.5\text{ }\mu\text{m}$  imaging from *Spitzer*/IRAC (Ashby et al. 2013; Steinhardt et al. 2014) to produce a final twelve-band catalogue. The process by which the catalogue is compiled is described in detail in sections 2.1, 2.3 and 2.5 of Mortlock et al. (2017). Table 1 gives the mean depths for each of the twelve bands across the wide and deep regions, which respectively make up 60% and 40% of the  $1\text{ deg}^2$  overlap region.

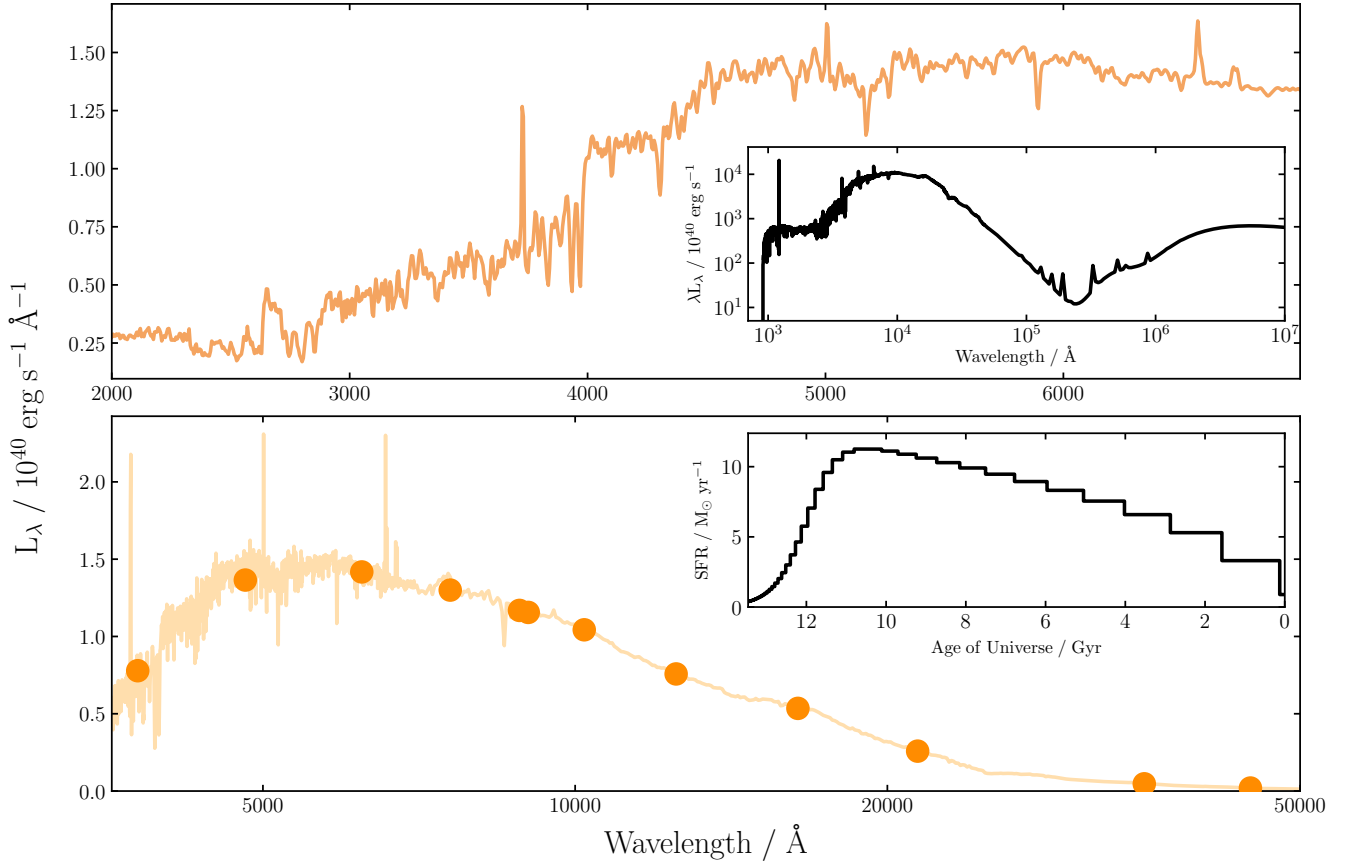
Mortlock et al. (2017) also calculate highly robust photometric redshifts for their whole catalogue using the median values from five different photometric-redshift codes. These median photometric redshifts have  $\sigma_{dz} \approx 0.02$ , where  $dz = (z_{\text{spec}} - z_{\text{phot}})/(1 + z_{\text{spec}})$ , and a catastrophic outlier fraction ( $|dz| > 0.15$ ) of  $\sim 1\%$  (see their section 3 for more details). They also calculate mass-completeness limits as a function of redshift and clean the catalogue by performing star-galaxy separation, and by matching to radio and X-ray datasets to remove AGN contaminants. The combination of all of these high-quality datasets provides us with an extremely clean catalogue of unrivalled scope and depth for studying galaxy evolution.

## 3 THE BAGPIPES CODE

Bayesian Analysis of Galaxies for Physical Inference and Parameter ESTimation, or BAGPIPES, is a Bayesian spectral fitting code, designed to model the emission from galaxies from the far-ultraviolet to the microwave regimes, and to fit these models to arbitrary combinations of spectroscopic and photometric observational data using the MULTINEST nested sampling algorithm (Feroz & Hobson 2008; Feroz et al. 2009, 2013). BAGPIPES is written purely in the PYTHON programming language and considerable effort has been made to make the API as intuitive and user-friendly as possible. This section describes the BAGPIPES code. In Section 3.1 we describe how models are generated by the code, then in Section 3.2 we describe the fitting of these models to observational data. BAGPIPES is publicly available and fully documented at <https://bagpipes.readthedocs.io>.

**Table 2.** List of input parameters for the example BAGPIPES model shown in Fig. 1. All parameters are discussed in Section 3 except for  $\alpha$ ,  $\beta$  and  $\tau$  for the double-power-law SFH parameterisation, the functional form for which is given in Equation 10. Here we assume Solar metallicity to take the value  $Z_{\odot} = 0.02$ .

Global	Double-power-law	Dust	Nebular
$\sigma_{\text{vel}} = 300 \text{ km s}^{-1}$	$\log_{10}(M_{\text{formed}}/M_{\odot}) = 11$	Calzetti	$\log_{10}(U) = -3$
$a_{\text{BC}} = 0.01 \text{ Gyr}$	$Z = 0.8 Z_{\odot}$	$A_V = 0.2$	
$z = 0$	$\tau = 12 \text{ Gyr}$	$\epsilon = 3$	
	$\beta = 0.5$	$T = 30 \text{ K}$	
	$\alpha = 30$	$\beta = 1.5$	



**Figure 1.** An example of a model galaxy built with BAGPIPES. The parameters of the model are given in Table 2 and are designed to be representative of a present-day quenched galaxy. The top panel shows the spectral output, in this case over the range 2000–7000 Å with a sampling of 5 Å. The top inset panel shows the full spectrum for the model. The bottom panel shows the photometric output, in this case fluxes observed through the twelve filters of the UltraVISTA catalogue (orange circles) described in Section 2. The bottom inset panel shows the SFH of the galaxy, which has been modelled as a double-power-law.

### 3.1 Model generation

The first major aspect of BAGPIPES is its ability to rapidly generate physically realistic model galaxy spectra from the far-ultraviolet to the microwave regime. Models can be built up to the desired level of complexity by specifying a number of model components, for example dust and nebular-emission prescriptions and SFH components. BAGPIPES uses these components (in practice these are passed to the API as PYTHON dictionaries containing parameter values) to generate an internal model of a galaxy spectrum. From this model, the user can request spectral data covering a given wavelength range at a given sampling, photometric fluxes through a series of user-defined filters, and emission-line fluxes. An example model is shown in Fig. 1. The parameters of this model are listed in Table 2.

A model galaxy spectrum in BAGPIPES observed at redshift  $z_{\text{obs}}$  has a luminosity per unit rest-frame wavelength,  $\lambda$  of  $L_{\lambda}(\lambda)$ . The luminosity is constructed as a sum involving four ingredients:

- (i) Simple stellar-population models,  $\text{SSP}(a, \lambda, Z)$ , which are a function of  $\lambda$ , the age of the stellar population,  $a$ , stellar metallicity,  $Z$  and the initial mass function (described in Section 3.1.1).
- (ii) The star-formation history,  $\text{SFR}(t)$ , which is composed of a sum over one or more SFH components (described in Section 3.1.2).
- (iii) The transmission function of the ionized ISM,  $T^{+}(a, \lambda)$ , as defined by Charlot & Longhetti (2001), including absorption, line emission, ionized continuum emission and emission from warm dust within H II regions (described in Section 3.1.3).



(iv) The transmission function of the neutral ISM,  $T^0(a, \lambda)$ , due to diffuse dust attenuation and emission (described in Section 3.1.4).

These four ingredients are described in more depth in the following four sections. The ingredients are then combined to give the galaxy luminosity according to the relationship

$$L_\lambda(\lambda) = \sum_{j=1}^{N_c} \sum_{i=1}^{N_a} \text{SFR}_j(t_i) \text{SSP}(a_i, \lambda, Z_j) T^+(a_i, \lambda) T^0(a_i, \lambda) \Delta a_i \quad (1)$$

where  $i$  runs across the age bins used in BAGPIPES (see Section 3.1.1),  $\Delta a_i$  are the widths of these bins,  $j$  runs across SFH components (see Section 3.1.2),  $N_c$  is the number of SFH components and  $N_a$  the number of age bins. The distinction between times and ages is explained at the end of Section 1; here  $t_i = t(z_{\text{obs}}) - a_i$ .

Once the galaxy luminosity has been calculated, it is redshifted and converted into an observed frame flux density,  $f_{\lambda_{\text{obs}}}(\lambda_{\text{obs}})$ , where  $\lambda_{\text{obs}} = (1 + z_{\text{obs}})\lambda$ , using the relationship

$$f_{\lambda_{\text{obs}}}(\lambda_{\text{obs}}) = \frac{L_\lambda(\lambda)}{4\pi D_L(z_{\text{obs}})^2 (1 + z_{\text{obs}})} T_{\text{IGM}}(\lambda, z_{\text{obs}}), \quad (2)$$

where  $D_L(z_{\text{obs}})$  is the luminosity distance to redshift  $z_{\text{obs}}$  and  $T_{\text{IGM}}(\lambda, z_{\text{obs}})$  is the transmission function of the IGM, described in Section 3.1.5. Finally, if an output spectrum is requested (as opposed to photometry or emission-line fluxes),  $L_\lambda(\lambda)$  can be convolved with a Gaussian kernel (before applying Equation 2) to model the effects of velocity dispersion, as described in Section 3.1.6.

### 3.1.1 Stellar population synthesis

Stellar Population Synthesis (SPS) is not implemented directly in BAGPIPES. Instead, the code is designed to accept pre-defined SPS models in the form of grids of simple stellar-population (SSP) models of different ages across a range of metallicities.

The SPS models currently implemented in the code are the 2016 version of the [Bruzual & Charlot \(2003\)](#) (hereafter BC03) models. These differ from earlier versions by their use of the Medium-resolution Isaac Newton Telescope library of empirical spectra (MILES; [Falc3n-Barroso et al. 2011](#)) in the UV-optical spectral region. The models implemented within the code are constructed using a [Kroupa & Boily \(2002\)](#) initial mass function (IMF). BAGPIPES does not currently include the option to vary element abundance patterns (e.g. alpha enhancement), with the model set currently implemented having scaled Solar abundances. In the future we also hope to distribute the code with the Binary Population and Stellar Synthesis (BPASS, [Eldridge & Stanway 2009](#)) models.

When a set of SPS models is loaded by BAGPIPES, the models are resampled in age (using a weighted summation method) onto a grid of ages,  $a_i$ , from  $\log_{10}(a_i/\text{Gyr}) = 6.0$  to  $10.2$  with uniform width in  $\log_{10}(\Delta a_i/\text{Gyr})$ . The default spacing is  $0.1$  dex, which sets the value of  $N_a$  in Equation 1 to 43.

### 3.1.2 Star-formation histories

Star-formation histories in BAGPIPES are constructed from one or more components,  $j$ , each of which specifies some functional form for star-formation rate as a function of time,  $\text{SFR}_j(t)$ . The total SFR is given by summing over these components, i.e.

$$\text{SFR}(t) = \sum_{j=1}^{N_c} \text{SFR}_j(t). \quad (3)$$

When  $L_\lambda(\lambda)$  is calculated using Equation 1, the  $\text{SFR}_j(t)$  are evaluated for all  $t_i = t(z_{\text{obs}}) - a_i$  where  $a_i$  is less than  $t(z_{\text{obs}})$ , i.e. the age of the stellar population is less than the age of the Universe at  $z_{\text{obs}}$ . We set  $\text{SFR}_j(t_i)$  to zero for all  $t_i$  corresponding to ages greater than the age of the Universe.

One may specify an unlimited number of SFH components, each with an individual functional form (see Section 4.1 for further discussion). The options currently implemented are:

- Delta function
- Constant
- Exponentially declining (Equation 9)
- Delayed exponentially declining
- Log-normal
- Double-power-law (Equation 10)
- Custom (directly input an array of SFR values).

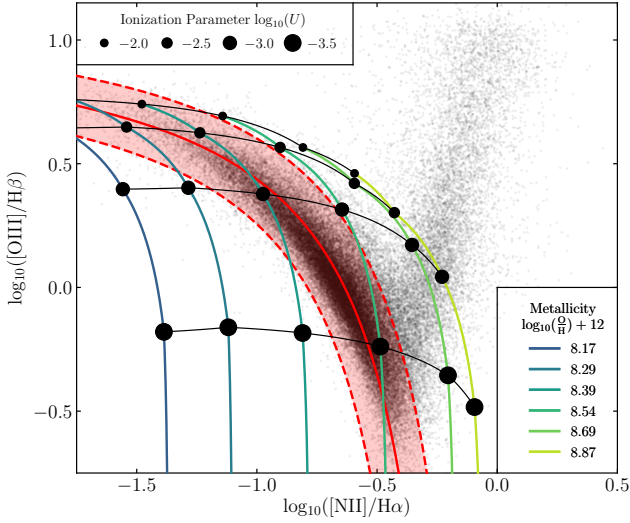
In this way, BAGPIPES can be used to generate a huge parameter space of possible SFHs, encompassing, for example, the non-parametric SFHs used by [Leja et al. \(2017\)](#) by the use of multiple constant components, or SFHs drawn from simulations of galaxy formation as in [Pacifci et al. \(2016\)](#). SFHs from simulations may be inputted either by assigning a burst component to each star particle (as in Section 4.3), or by loading tabulated SFHs from simulations directly into the code as custom SFH components.

Each component also requires the specification of a total mass of stars formed,  $M_{\text{formed}}$ , by that component over its whole history, and a metallicity value,  $Z_j$ , which is generated by linearly interpolating the  $\text{SSP}(a_i, \lambda, Z_j)$  between different grids of models. Thus, metallicity evolution can be modelled by specifying different metallicities for multiple SFH components covering different epochs of cosmic time. BAGPIPES also includes basic functionality for specifying a distribution function for the metallicities of stars in a galaxy, as in [Leja et al. \(2017\)](#), as an alternative to linear interpolation between grids. In the future we will extend this scheme to allow full chemical-evolution histories to be specified, permitting metallicity distribution functions which are also a function of cosmic time.

### 3.1.3 Nebular emission

The nebular-emission model implemented in BAGPIPES is constructed following the methodology of [Byler et al. \(2017\)](#), using the latest (2017) version of the CLOUDY photoionization code ([Ferland et al. 2017](#)). We model H II regions with a spherical shell geometry of fixed radius and assume the nebular emission from a galaxy is the sum of emission from H II regions of different ages, as in [Charlot & Longhetti \(2001\)](#). We assume the metallicity of the ionized gas is the same as the stars which produce the ionizing photons and use the Solar abundances of [Anders & Grevesse \(1989\)](#) and ISM depletion factors and Helium and Nitrogen scaling relations of [Dopita et al. \(2000\)](#). All of our models include dust grains using the ‘‘ISM’’ prescription within CLOUDY, which has a grain size distribution and abundance pattern designed to reproduce the observed extinction properties for the ISM of the Milky Way.

BAGPIPES includes functionality for computing grids of nebular-emission models corresponding to grids of input SSP models (these have been pre-computed for BC03). CLOUDY is run using each SSP in turn as the input spectrum whilst varying the logarithm of the ionization parameter,  $U$ , in steps of  $0.5$  between  $\log_{10}(U) = -4$  and  $-2$  by varying the number of hydrogen-ionizing photons,  $Q_{\text{H}}$  at a fixed Hydrogen density of  $100 \text{ atoms cm}^{-3}$ . At each value of  $\log_{10}(U)$ ,  $a_i$  and metallicity we save the output diffuse continuum,



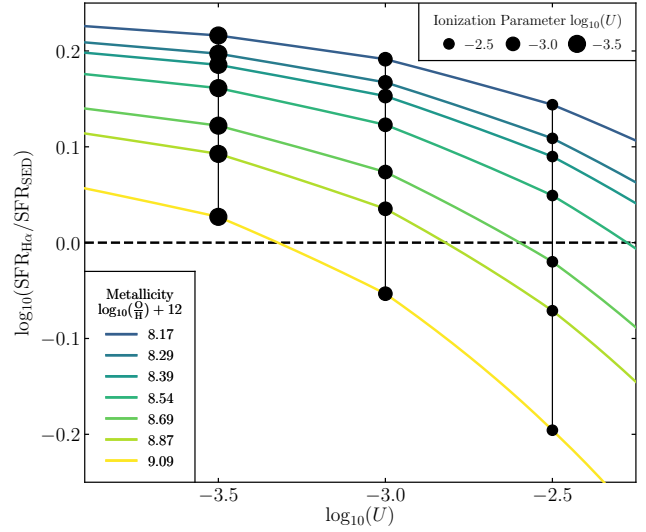
**Figure 2.** BPT diagram with the positions of BAGPIPES models overlaid as a function of metallicity and ionization parameter. Black lines are lines of constant ionization parameter and coloured lines are lines of constant metallicity. The theoretically calibrated star-forming sequence for galaxies at  $z \sim 0$  of Kewley et al. (2013) is shown in red and a sample of local SDSS galaxies from Brinchmann et al. (2013) is shown in grey. The BAGPIPES models are BC03-based, with constant SFR over the last 10 Myr. Our models show good overall agreement with the models of Kewley et al. (2013) and the results of Moustakas et al. (2006), who find local star-forming galaxies fall within the range  $-3.8 < \log_{10}(U) < -2.9$  and  $8.15 < 12 + \log(\frac{O}{H}) < 8.7$ .

which includes contributions from ionized gas and warm dust, and fluxes for a series of emission lines.

The base list of lines we employ is that given in table 3 of Byler et al. (2017), containing 128 separate features. The features have been renamed in the new version of CLOUDY, due to slight shifts in wavelength, so it was necessary to manually match-up new and old labels. In addition, the five narrowly spaced C II lines around 2326 Å were replaced with the total flux for a blend of these lines, as identified by ‘Blnd 2326.00A’, and the two narrowly spaced He I lines around 1.083 μm were replaced with ‘TOTL 1.08303m’. Finally we added the He II feature at 4686 Å, denoted in CLOUDY by ‘He 2 4685.64A’ to our list, meaning that, in total, we track 124 separate emission features.

When including nebular emission in a BAGPIPES model, one must specify the ionization parameter and lifetime of H II regions (or stellar birth-clouds),  $a_{BC}$ . The nebular-continuum and line models corresponding to all  $SSP(a_i, \lambda, Z_j)$  with  $a_i < a_{BC}$  are then linearly interpolated on-the-fly in  $\log_{10}(U)$  and metallicity, and added to the corresponding  $SSP(a_i, \lambda, Z_j)$  model. Contrary to Equation 1 this is done by addition to, rather than multiplication of  $SSP(a_i, \lambda, Z_j)$ , however the effect is the same as setting  $T^+(a_i, \lambda) > 1$ .

At this stage, the total energy in the combined (emission lines plus continuum) nebular model is set to the same as the total energy in the hydrogen-ionizing photons from all SSPs with  $a_i < a_{BC}$ . The hydrogen-ionizing continuum is then removed from these SSPs (we set  $T^+(a_i, \lambda) = 0$  for  $a_i < a_{BC}$  and  $\lambda < 911.8$  Å; effectively assuming an escape fraction of zero) in order to maintain energy conservation. As well as being inserted into the output spectrum, observed fluxes for each emission feature are propagated through BAGPIPES separately, meaning line fluxes can be accessed directly by the user.



**Figure 3.** The offset between H $\alpha$ -derived SFR for BAGPIPES models using the Kennicutt & Evans (2012) calibration and input SFR as a function of metallicity and ionization parameter. We find a slight excess of H $\alpha$  flux in our models compared to those predicted by Kennicutt & Evans (2012), except at supersolar metallicities or log-ionization parameters above  $-2.5$ . This is likely due to our assumption of an escape fraction of zero. Our models are consistent to within the observed scatter on the Kennicutt & Evans (2012) relationship.

We perform two consistency checks on our grid of BC03-based nebular-emission models to demonstrate that they are in agreement with similar predictions from the literature. We firstly check the optical line ratios as a function of metallicity and ionization parameter by plotting our models on the Baldwin-Phillips-Terlevich (BPT) diagram (Baldwin et al. 1981) in Fig. 2. We confirm our predictions are consistent with those of Kewley et al. (2013) by comparison with their figure 1. We also demonstrate that the range of metallicities and ionization parameters measured by Moustakas et al. (2006) for local galaxies would be correctly reproduced by our models.

Having confirmed this, we then check that the line strengths are realistic in Fig. 3 by comparing our input star-formation rate,  $SFR_{SED}$  with those derived from the strengths of our H $\alpha$  lines using the calibration of Kennicutt & Evans (2012),  $SFR_{H\alpha}$ . We find that our H $\alpha$  fluxes tend to be slightly higher than those predicted by Kennicutt & Evans (2012), e.g. by 0.08 to 0.19 dex for metallicities of  $12 + \log(\frac{O}{H}) = 8.69$  and  $8.17$  respectively at  $\log_{10}(U) = -3.0$ , which is consistent to within the observed scatter on the relationship. This overestimation is likely due to our assumption of an escape fraction of zero; an escape fraction of  $\sim 20$  per cent would be required to bring our  $12 + \log(\frac{O}{H}) = 8.69$ ,  $\log_{10}(U) = -3.0$  result into line with Kennicutt & Evans (2012).

### 3.1.4 Dust attenuation and emission

Dust attenuation in BAGPIPES is designed in a modular fashion such that different dust curves may be easily implemented. Each dust-attenuation curve takes the overall functional form

$$\log_{10}(T^0(a_i, \lambda)) = \begin{cases} -\frac{0.4 \epsilon_{AV} k(\lambda)}{R_V} & a_i < a_{BC} \\ -\frac{0.4 A_V k(\lambda)}{R_V} & a_i > a_{BC} \end{cases} \quad (4)$$

where  $a_{\text{BC}}$  is the lifetime of stellar birth-clouds as described in Section 3.1.3,  $A_V$  is the attenuation in the V band ( $\sim 5500 \text{ \AA}$ ) in magnitudes,  $\epsilon$  is a constant which can be used to control the extra attenuation towards H II regions and  $k(\lambda)$  and  $R_V$  are specific to the dust model being used.

Three dust models are currently implemented: the Calzetti et al. (2000) law for local star-forming galaxies, the Cardelli et al. (1989) Milky Way dust law, and a flexible model based on that of Charlot & Fall (2000). For the first two of these models the form of  $k(\lambda)$  and the value of  $R_V$  are specified in the paper cited. For use in BAGPIPES, these are extrapolated to shorter wavelengths using a power-law fit to the dust curve in the near-ultraviolet. For the Charlot & Fall (2000) model we define

$$\frac{k(\lambda)}{R_V} = (\lambda/5500\text{\AA})^{-n} \quad (5)$$

where  $n$  is the slope of the attenuation law. Our  $\epsilon$  parameter is the reciprocal of the parameter  $\mu$  often used in the literature (e.g. Cullen et al. 2017b) for models of this type.

Dust emission from the neutral ISM is currently modelled as a single-temperature greybody, with flux per unit frequency,  $S_{\text{gb}}(\nu)$  (e.g. Hildebrand 1983; Younger et al. 2009) given by

$$S_{\text{gb}}(\nu) \propto \nu^\beta B_\nu(T) = \frac{2h}{c^2} \frac{\nu^{\beta+3}}{e^{\frac{hc}{\lambda kT}} - 1} \quad (6)$$

where  $T$  is the temperature of the greybody,  $\beta$  is the spectral emissivity index and the dust has been assumed to be optically thin. The normalisation of this component is set such that the total energy removed from the spectrum by dust attenuation is the same as the total energy emitted by the greybody.

Dust emission in BAGPIPES is hence modelled by two separate components. Firstly, the hot-dust component included in the CLOUDY output diffuse continua for H II regions (described in Section 3.1.3), and secondly the greybody component due to cold, diffuse dust emission. This kind of two-component approach has been shown to be successful in modelling observed infrared SEDs (e.g. Casey 2012), and work is ongoing to test our model against observational data at these wavelengths. In the future we hope to implement the option of a more complex physical model for dust emission.

### 3.1.5 IGM attenuation

For  $T_{\text{IGM}}(\lambda, z_{\text{obs}})$ , BAGPIPES incorporates the IGM attenuation model of Inoue et al. (2014), an updated version of the Madau (1995) model. The analytic expression presented in their section 4 is calculated and tabulated for rest-frame wavelengths between 911.8 Å and 1215.7 Å, then interpolated by the code for use in model generation. We assume  $T_{\text{IGM}}(\lambda, z_{\text{obs}}) = 0$  for any flux at  $\lambda < 911.8 \text{ \AA}$ , the majority of which is already removed when applying the nebular-emission prescription (see Section 3.1.3).

### 3.1.6 Velocity dispersion and spectral sampling

BAGPIPES is designed to perform spectroscopic fitting as well as fitting to photometry, meaning the effects of velocity dispersion,  $\sigma_{\text{vel}}$ , must be included in the model to match observed spectral features. To facilitate this, BAGPIPES converts the wavelength sampling of all input SPS models to constant spectral resolution,  $R = \frac{\lambda}{\Delta\lambda}$ . By default, a low resolution of  $R = 100$  is used over regions which

will be used only for calculation of output photometry, and a higher resolution of  $R = 600$  is used over regions which will be used to generate the output spectrum (in practice,  $R$  is doubled to achieve Nyquist sampling at the resolutions quoted). The latter region is then convolved with a Gaussian kernel in velocity space to model the effect of velocity dispersion within the observed galaxy, which is assumed to be the same for the stellar and gas components.

## 3.2 Model fitting

The second major aspect of BAGPIPES is its ability to fit the models described in Section 3.1 to observational data within the framework of Bayesian inference. These data can take the form of an observed spectrum and/or any number of photometry points, all of which can be fitted simultaneously by the code. In Section 3.2.1 we outline the principles of our fitting method as applied to photometric observations, additional considerations when fitting spectroscopic observations will be addressed in a follow-up paper. In Section 3.2.2 we outline how fitting is performed using the MULTINEST nested sampling algorithm. Section 3.2.3 is a brief note on specifying prior distributions within BAGPIPES.

### 3.2.1 Bayesian inference methodology

Bayesian inference methods use Bayes' theorem to update prior knowledge about the probability of a hypothesis, or model, based on new data. This can take the form of parameter estimation, in which constraints on the parameters of a given model are updated, or model selection, in which the relative probabilities of two or more different models are assessed. Bayes' theorem states that for some new data,  $\mathcal{D}$  and a hypothesis,  $\mathcal{H}$  with parameter vector  $\Theta$ ,

$$P(\Theta | \mathcal{D}, \mathcal{H}) = \frac{P(\mathcal{D} | \Theta, \mathcal{H}) P(\Theta | \mathcal{H})}{P(\mathcal{D} | \mathcal{H})}. \quad (7)$$

Here,  $P(\Theta | \mathcal{H})$  is the prior probability distribution for the model parameters: what is known before the new data are considered.  $P(\mathcal{D} | \Theta, \mathcal{H}) = \mathcal{L}(\mathcal{D} | \Theta, \mathcal{H})$  is the likelihood: the probability of obtaining the new data under the assumption of a certain set of parameter values.  $P(\Theta | \mathcal{D}, \mathcal{H})$  is the posterior probability distribution: what is known once the new data have been considered. Finally,  $P(\mathcal{D} | \mathcal{H})$  is the Bayesian evidence, or marginal likelihood: how good the hypothesis is overall at explaining the data.

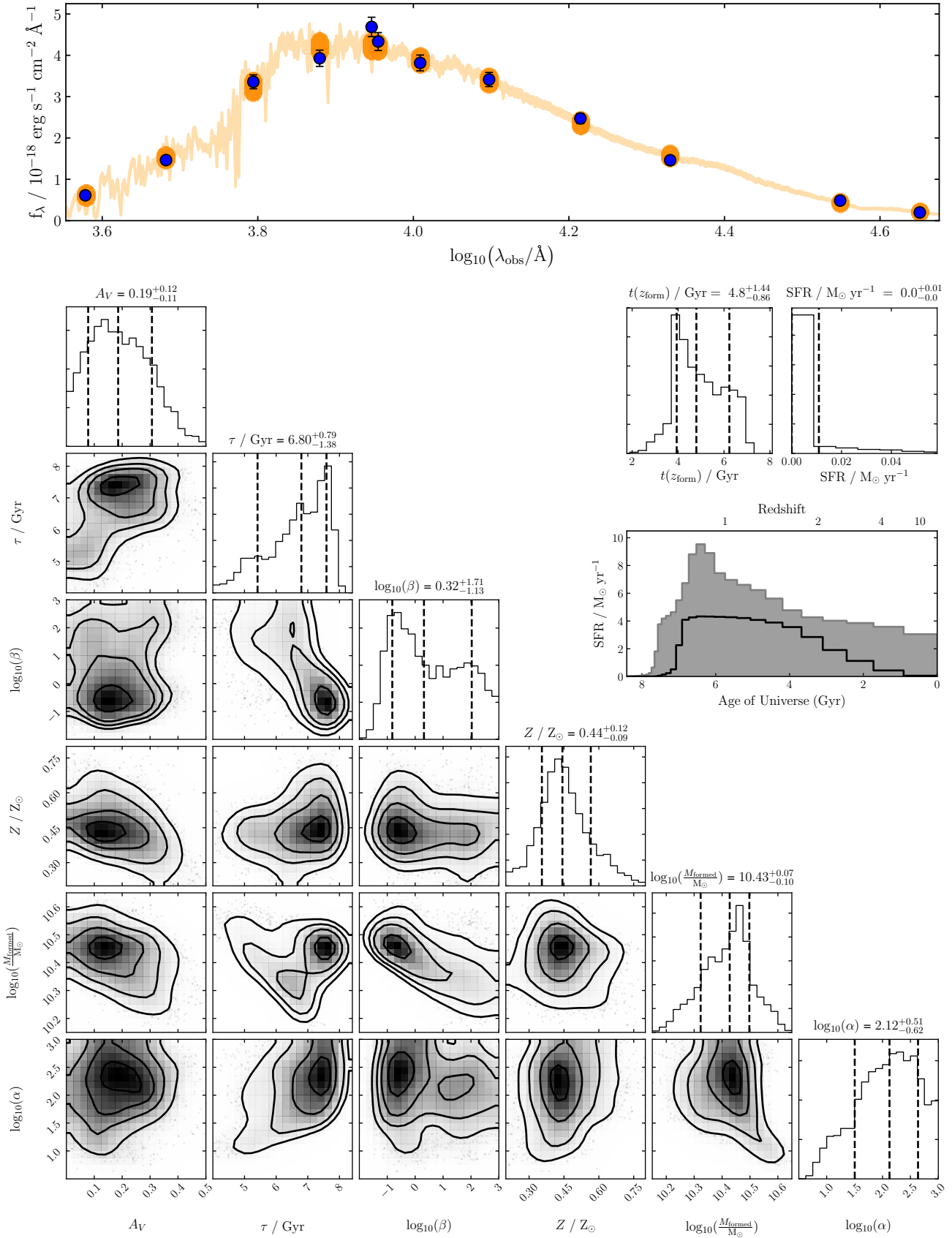
For our purposes, the hypothesis consists of a certain model parameterisation within BAGPIPES (e.g. type of dust model, number and type of SFH components, whether nebular emission is included etc). Once the model has been constructed, the user then defines prior probability distributions for its parameters (see Section 3.2.3 for details), and provides observational data,  $f_i$  with associated uncertainties  $\sigma_i$ , which will be used to constrain those parameters.

The observational data are incorporated into the calculation through the likelihood function, which is constructed assuming that uncertainties are Gaussian and independent. When fitting only photometric data, the log-likelihood function used by BAGPIPES is given by (e.g. Hogg et al. 2010)

$$\ln(\mathcal{L}) = -0.5 \sum_i \ln(2\pi\sigma_i^2) - 0.5 \sum_i \frac{(f_i - f_i^{\mathcal{H}}(\Theta))^2}{\sigma_i^2} \quad (8)$$

where  $f_i^{\mathcal{H}}(\Theta)$  is the model prediction corresponding to the observed flux  $f_i$ , which is obtained by generating a BAGPIPES model as described in Section 3.1.





**Figure 4.** Example of BAGPIPES output when fitting mock photometry from the MUFASA simulation for an object at  $z_{\text{obs}} = 0.5$  using a double-power-law SFH. The mock photometry is shown on the top panel in blue (the bands are those of Table 1). The 16th to 84th percentile range for the posterior spectrum and photometry are shaded orange. A corner plot showing the posterior for fitted parameters is shown below, with posterior SFH information shown to the right.

### 3.2.2 Nested Sampling implementation

BAGPIPES makes use of the nested sampling algorithm of Skilling (2006) to obtain the posterior distribution and evidence value given some model, prior distribution and observational data. Nested sampling is implemented in BAGPIPES using MULTINEST (Feroz & Hobson 2008; Feroz et al. 2009, 2013), accessed through the PYMULTINEST interface (Buchner et al. 2014).

Nested sampling allows for efficient exploration of higher-dimensional, multimodal and highly degenerate parameter spaces. This is invaluable in many circumstances relevant to spectral fitting, notably when dealing solely with broad-band photometric observations. In this case, the age-metallicity-dust degeneracy often leads to, at best, poorly constrained and highly degenerate parameter estimates, or at worst several widely spaced local minima in parameter space (e.g. McLure et al. 2011), any of which can trap traditional numerical functional minimisation routines or MCMC methods.

Once the posterior distribution is returned by MULTINEST, it can be post-processed by BAGPIPES to obtain other posterior information of interest, such as the posterior spectrum and SFH. A variety of visualisations, such as corner plots (Foreman-Mackey 2016) can also be constructed. An example corner plot and posterior spectral plot from a fit to mock photometry for one of the MUFASA objects introduced in Section 4 is shown in Fig. 4.

In the future, we intend to diversify the fitting options available in BAGPIPES, giving the user access to more advanced versions of the nested sampling algorithm. This is necessary for faster and more reliable sampling of parameter spaces with higher dimensionalities, such as those encountered when fitting spectroscopic data. In particular we hope to include the option to implement dynamic nested sampling through the DYNesty package (Speagle et al. in prep).

### 3.2.3 Specifying prior distributions on model parameters

Any parameter which can be specified in the construction of a BAGPIPES model, as described in Section 3.1, can be fitted using the code. When fitting a parameter, a prior probability distribution must be specified, consisting of an upper and lower limit on the parameter value, and a functional form for the prior probability density between these limits. In certain cases, these probability density functions will include hyperparameters which must also be specified. BAGPIPES includes a separate “priors” module which allows the user to specify their own prior distributions, or pick from a number of options which have already been implemented.

## 4 TESTING STAR-FORMATION HISTORY MODELS WITH MUFASA

BAGPIPES is a powerful tool for studying galaxy evolution, however, before applying it to real data it is important to test whether the models we define within it produce unbiased estimates of the parameters of interest. For example, it is often possible to obtain a good fit to photometric observations of a galaxy by modelling the whole SFH as a single burst of star formation (SSP model). This is likely to produce a reasonably accurate and unbiased estimate of, for example, the redshift of the galaxy. However, if this method were used to estimate the average time at which stars in the galaxy formed (see Section 4.2), it would clearly be biased towards younger ages, as the luminosities of galaxies are dominated by the youngest stars they contain. This is an example of model misspecification, where the chosen functional form for a model, or aspect of a model (in this

case the SFH), is adequate for obtaining a fit to the data, however fails to accurately represent the underlying data-generating process, leading to biases in derived parameters.

Recently, Leja et al. (2017) performed tests to validate the PROSPECTOR inference framework by comparing physical properties derived from fitting only broad-band photometry against well-calibrated indicators derived from aperture-matched spectroscopy for a sample of local galaxies. They show that their model is capable of unbiased estimation of the SFR, dust-attenuation properties and stellar metallicities of their galaxies.

In this work, however, our focus is on the SFH properties of galaxies. This presents an additional challenge, as no independent, well-calibrated methods exist to obtain, for example, the average time at which the stars in a galaxy formed. Instead we look to simulations of galaxy formation to provide realistic SFHs which can be converted into mock photometric observations. These can then be fitted with BAGPIPES using a variety of different model parameterisations with different priors, in order to find a model which accurately reproduces the properties of the input SFHs.

In this section we use the MUFASA suite of cosmological hydrodynamic simulations (see Section 4.3) to construct a catalogue of mock observations for simulated quenched galaxies with realistic SFHs, chemical-enrichment histories and dust properties. We tailor our mock observations to match as closely as possible the UltraVISTA data introduced in Section 2. We then test the ability of a variety of models within BAGPIPES to reproduce the SFH properties of quenched MUFASA galaxies.

We begin in Section 4.1 by discussing the advantages and disadvantages of different commonly used representations of the SFHs of galaxies. In Section 4.2, we introduce a set of parameters which can be used to compare how well different SFH models reproduce the important features of the input MUFASA SFHs. These parameters will also be used in the presentation of our results in Section 6. We then give details of the MUFASA simulation and the process of generating mock observations in Section 4.3. Finally, in Section 4.4 we fit these mock observations using BAGPIPES, and compare the abilities of the exponentially declining and double-power-law SFH parameterisations to recover the SFH properties of MUFASA quenched galaxies. We demonstrate that, by using a double-power-law SFH model with the correct priors (see Table 3), we can recover unbiased estimates of the underlying properties of the MUFASA SFHs. In contrast, we show that the exponentially declining model returns small but significant biases in these parameters. Based on these results, we proceed to use our double-power-law SFH model to fit the UltraVISTA dataset in Section 5.

### 4.1 Common SFH parameterisations

The traditional approach to SED fitting has been to use a simple functional form to parameterise the SFH, most commonly the exponentially declining function

$$\text{SFR}(t) \propto \begin{cases} e^{-\frac{t-T_0}{\tau}} & t > T_0 \\ 0 & t < T_0 \end{cases} \quad (9)$$

where  $T_0$  and  $\tau$  are parameters to be fitted. The main advantage of this rigid parameterisation is the speed of fitting, however Maraston et al. (2010) and Reddy et al. (2012) have shown by that it becomes less effective at higher redshifts, and Wuyts et al. (2011) and Pforr et al. (2012) have shown that the biases on the estimated SFH parameters (e.g. the ongoing SFR) are highly dependent on the permitted ranges of (alternatively, the priors on) model parameters.

It is possible to define more flexible parameterisations with the same number of parameters, such as the log-normal form of [Gladsters et al. \(2013\)](#), which has been shown by [Diemer et al. \(2017\)](#) to produce good fits to SFHs from the cosmological simulation Illustris ([Vogelsberger et al. 2014](#)). This form is clearly more physical, as star formation is required to increase smoothly from zero at the beginning of the Universe, rather than jumping from zero to its maximum value several billion years later. However, when using the log-normal form, star formation must always decline more slowly than it rises, meaning that this parameterisation struggles to model rapidly quenching systems.

Attempts have also been made to improve the flexibility of model SFHs by using more complex parameterisations, at the cost of increased computational expense. One option is the double-power-law form, which has previously been used to fit the evolution of the cosmic SFR density (e.g. [Behroozi et al. 2013](#)),

$$\text{SFR}(t) \propto \left[ \left( \frac{t}{\tau} \right)^\alpha + \left( \frac{t}{\tau} \right)^{-\beta} \right]^{-1} \quad (10)$$

where  $\alpha$  and  $\beta$  are the falling and rising slopes respectively, and  $\tau$  is related to the time at which star formation peaks. The major advantage of the double-power-law SFH model is the decoupling of the rising and falling slopes of the SFH, allowing, in contrast to the log-normal form, a slow rise and rapid cutoff to star formation, as demonstrated in Fig. 1. The double-power-law form is also discussed by [Diemer et al. \(2017\)](#) as a possible improvement on their log-normal parameterisation. They confirm that the double-power-law parameterisation produces better agreement with Illustris, particularly for rapidly quenching galaxies.

We also note that progress has been made in SED fitting using libraries of simulated SFHs from cosmological simulations (e.g. [Pacifci et al. 2012](#)), and using non-parametric SFHs (see Section 3.1.2). However these approaches are significantly more computationally intensive due to the inclusion of many free parameters, and are hence best suited to smaller sample sizes, rather than the large photometric catalogue we consider in this work. Testing of these methods within BAGPIPES for smaller datasets is currently ongoing.

## 4.2 Comparisons between SFH parameterisations

When attempting to compare results obtained using different SFH parameterisations, or to compare fitted SFHs with the input SFHs from MUFASA, it is necessary to define common parameters which can be derived from any SFH. One such parameter which is readily available is the living stellar mass present (excluding remnants) at the redshift of observation,  $M_*(z_{\text{obs}})$ , or simply  $M_*$ .

To answer Question (i), posed in Section 1, we require a method of quantifying when this mass was assembled. One common approach (e.g. [Thomas et al. 2017](#)) is to calculate the mass-weighted age, or in this case the mass-weighted time (measured forwards from the beginning of the Universe; see the note at the end of Section 1) of a galaxy,  $t_{\text{MW}}$ , given by

$$t_{\text{MW}} = \frac{\int_0^{t_{\text{obs}}} t \text{SFR}(t) dt}{\int_0^{t_{\text{obs}}} \text{SFR}(t) dt} = t(z_{\text{form}}) \quad (11)$$

where  $t_{\text{obs}} = t(z_{\text{obs}})$ . We then define the redshift of formation,  $z_{\text{form}}$ , based on this mass-weighted time by setting  $t_{\text{MW}} = t(z_{\text{form}})$ .

In order to answer Question (ii), we require knowledge of the redshifts at which galaxies quenched,  $z_{\text{quench}}$ . In order to define this

we first require a definition of a quenched galaxy. This will also be necessary for selecting our quenched sample from the UltraVISTA catalogue. Two common methods exist for selecting samples of quenched galaxies, firstly based on their rest-frame UVJ colours ([Williams et al. 2009](#)), and secondly using a cut in specific SFR (sSFR; SFR divided by  $M_*$ ), often evolving with observed redshift (e.g. [Pacifci et al. 2016](#)).

Relating the SFH of a galaxy to its rest-frame UVJ colours at earlier times requires several assumptions, such as the time-evolution of dust attenuation. This makes generalising UVJ selection to the past SFHs of galaxies extremely challenging. We therefore use a method similar to a specific SFR cut, by defining the dimensionless normalised SFR (nSFR), which is the current SFR as a fraction of the time-averaged SFR over the whole SFH, given by

$$\text{nSFR}(t) = \frac{\text{SFR}(t)}{\langle \text{SFR} \rangle} = \text{SFR}(t) \frac{t}{\int_0^t \text{SFR}(t') dt'} = \frac{t \text{SFR}(t)}{M_{\text{formed}}} \quad (12)$$

where  $\int_0^t \text{SFR}(t') dt' = M_{\text{formed}}$  is the total mass of stars formed by the galaxy up to the time  $t$ , which is closely related to the living stellar mass.

We define quenched galaxies as those which have normalised star-formation rate  $< 0.1$ , meaning the SFR at the time of observation must be less than 10 per cent of the average SFR across the history of the galaxy. This is effectively similar to an evolving cut in specific SFR, however it is not biased by the epoch at which the galaxy's stellar mass was assembled, as it depends on the total stellar mass formed rather than the living stellar mass. The value of 0.1 was chosen to preserve good agreement with UVJ selection (as demonstrated in Section 5). We show in Fig. 5 how our cut in normalised SFR corresponds to a cut in sSFR and confirm that our selection criterion is very similar to that of [Pacifci et al. \(2016\)](#), which they also find to be in good agreement with UVJ selection. We hence define  $z_{\text{quench}}$  as the redshift at which the normalised star-formation rate falls below 0.1. In the case of the MUFASA SFHs, which can be bursty and therefore fall below this threshold several times, we select the latest time at which this happens.

Finally, for our analysis in Section 6 it will be useful to define the quenching timescale,  $\Delta t_{\text{quench}}$  for our galaxies. We define this as being the duration between the time of formation  $t(z_{\text{form}})$ , or mass-weighted time, and the time of quenching,  $t(z_{\text{quench}})$ , such that

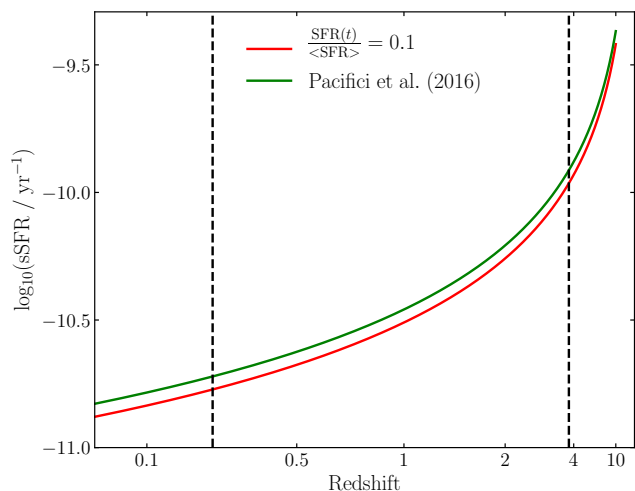
$$\Delta t_{\text{quench}} = t(z_{\text{quench}}) - t(z_{\text{form}}). \quad (13)$$

This definition of  $\Delta t_{\text{quench}}$  is designed to trace the timescale over which star-formation declines from its peak to a normalised SFR of 0.1, at which point we define the galaxy as quenched.

Figure 6 shows how successful  $\Delta t_{\text{quench}}$  is in capturing this for a range of different SFH shapes. As can be seen, a rapidly quenched galaxy which has an extended rising wing will have a higher  $\Delta t_{\text{quench}}$  than one which quenches equally quickly, but also forms very rapidly. Thus  $\Delta t_{\text{quench}}$  traces both the speed of quenching as intended, but also has a secondary dependence on how extended the SFH is prior to the onset of quenching.

As the dynamical timescales of galaxies evolve with redshift, a parameter of significant interest is  $\Delta t_{\text{quench}}$  as a fraction of the age of the Universe when quenching occurs,  $t(z_{\text{quench}})$ . This allows us to directly compare quenching timescales for galaxies at different observed redshifts. We define this normalised quenching timescale as  $\tau_{\text{quench}}$ , where

$$\tau_{\text{quench}} = \frac{\Delta t_{\text{shutoff}}}{t(z_{\text{quench}})}. \quad (14)$$



**Figure 5.** Selection criteria for quenched galaxies plotted as cuts in specific SFR as a function of redshift. Our cut in normalised SFR at 0.1 (see Section 4.2) is shown in red, assuming that the living stellar mass,  $M_*$  is 0.25 dex less than the total stellar mass formed (in reality this is dependent on the shape of the SFH). The selection criterion employed by Pacifci et al. (2016) is shown in green. These can be seen to be very similar, and both produce good agreement with the UVJ selection criteria of Williams et al. (2009). The black dashed lines denote the edges of the redshift range spanned by our UltraVISTA sample.

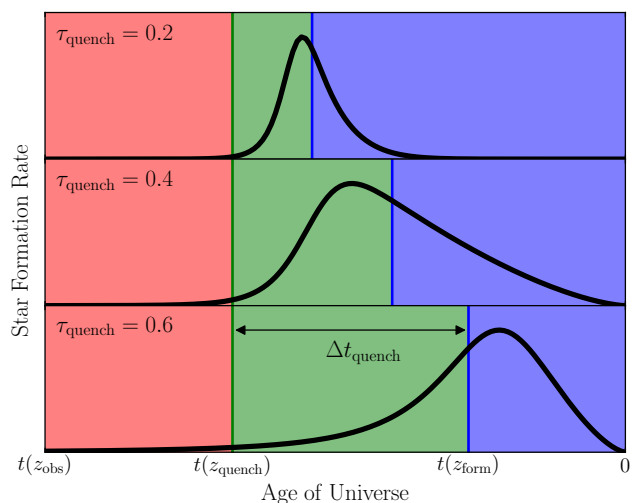
### 4.3 Generating mock observations of MUFASA galaxies

MUFASA (Davé et al. 2016) is a suite of cosmological hydrodynamic simulations which correctly reproduces many of the observed properties of the population of quenched galaxies at  $z = 0$ , including the slope of the red sequence and downsizing trend, as well as correct number densities for high-mass ( $M_* \gtrsim 10^{10.7} M_\odot$ ) quenched galaxies out to  $z \sim 2$  (Davé et al. 2017). This makes MUFASA an ideal resource for testing the ability of our model to recover realistic SFHs for massive quenched galaxies across a wide range of redshifts.

We use the post-processed data-products from MUFASA described by Davé et al. (2017), which consist of a list of star particles for each galaxy with individual masses, ages, metallicities and  $A_V$  values. We use snapshots of the simulation taken between  $z = 0.5$  and  $2.5$  at intervals of  $\Delta z = 0.5$ . To begin with, we select all galaxies from each of these snapshots with  $M_* > 10^{10} M_\odot$  in order to match the sample of observed galaxies we intend to extract from UltraVISTA (see Section 5).

For each MUFASA galaxy we generate a spectrum using BAGPIPES by modelling each star particle as a burst of star formation. We use the BC03 models and the Calzetti et al. (2000) attenuation curve, which has been shown recently to be a robust average of the attenuation curves observed for galaxies across a wide range of redshifts (e.g. Cullen et al. 2017a,b; McLure et al. 2017; but see Kriek & Conroy 2013). We also fix  $\log_{10}(U) = -3$  and  $a_{BC}$  to 10 Myr for our nebular-emission model, as these properties have only marginal impact on broad-band photometry for quiescent galaxies. For each galaxy we then calculate observed magnitudes in the twelve filters of the UltraVISTA catalogue described in Section 2.

We also compile the simulated SFHs of these objects in 100 Myr bins and calculate the normalised SFR (Equation 12) at the time of observation using the SFR in the most recent 100 Myr bin as the current ongoing SFR. We then select objects with normalised



**Figure 6.** Pictorial representation of the scheme for describing and comparing galaxy SFHs introduced in Section 4.2. The three example double-power-law SFHs shown all quench at the same time but have different shapes, leading to different values of  $\tau_{\text{quench}}$ , which can be seen to trace both the speed of quenching, and, to a lesser extent, how quickly the SFH rises. The quenching timescale can be thought of as the time between peak star formation and the time of quenching.

$\text{SFR} < 0.1$  as our quenched sample, as described in Section 4.2. For the SFHs of our quenched sample we then calculate the three indicators described in Section 4.2,  $M_*$ ,  $z_{\text{quench}}$  and  $z_{\text{form}}$ .

Once the sample had been selected, Gaussian noise was added to the mock photometry at the levels listed for the UltraVISTA wide strips in Table 1. An error floor was added such that no photometry point corresponded to a detection of greater than  $20\sigma$ , or, in the case of the IRAC channels,  $10\sigma$ . Finally a scatter was added to the catalogued observed redshifts of the MUFASA galaxies at the same level ( $dz = 0.02$ ) as the photometric redshifts calculated for our UltraVISTA catalogue. At the end of this process we had built a catalogue of mock observations for 677 simulated quenched galaxies, tailored to match as closely as possible the sample we intended to select from the UltraVISTA data of Section 2.

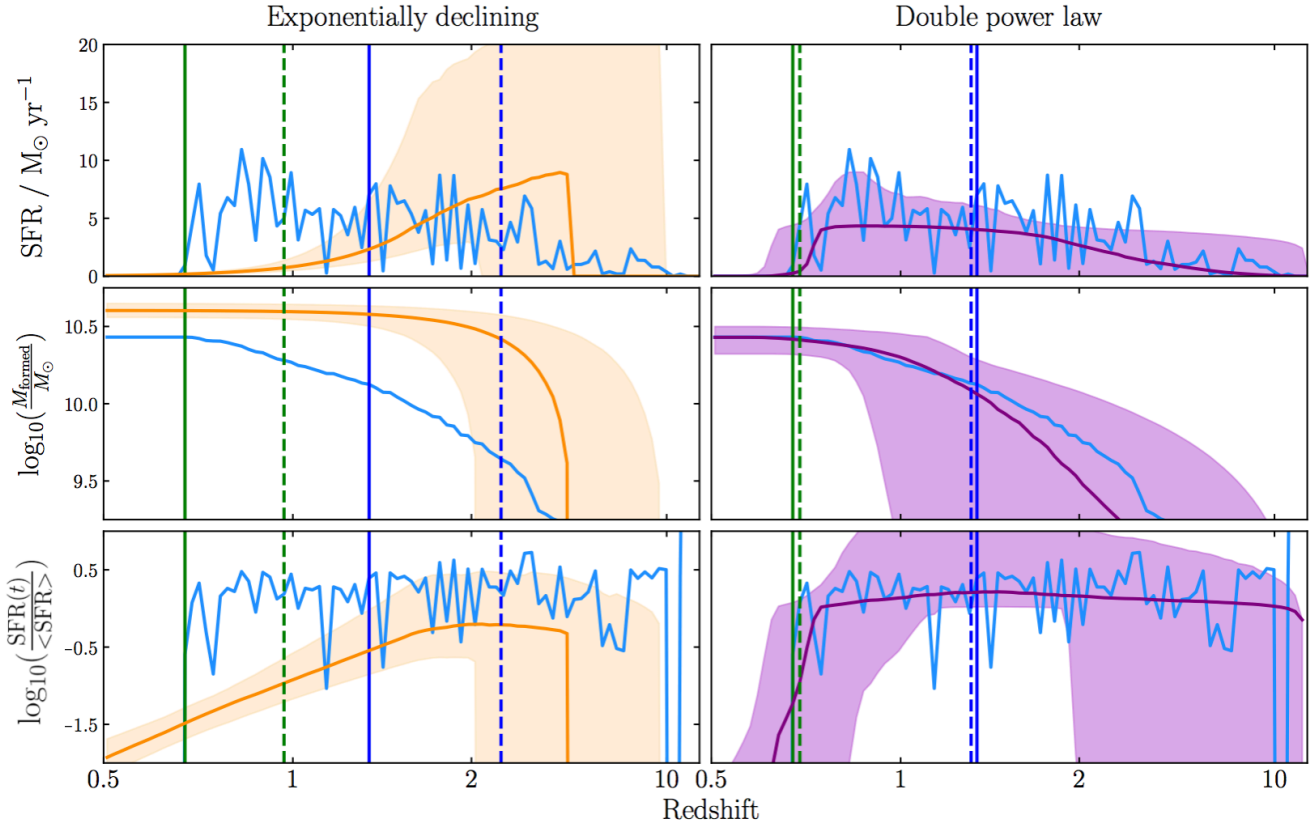
### 4.4 Recovering star-formation histories for MUFASA quenched galaxies

In this section we present the results of fitting the mock observations produced in Section 4.3 with different SFH parameterisations. We are interested primarily in SFH recovery, and hence leave discussion of other physical parameters to future work. We evaluate our ability to accurately recover the MUFASA SFHs using the three indicators introduced in Section 4.2,  $M_*$ ,  $z_{\text{quench}}$  and  $z_{\text{form}}$ .

For all of our fitted models we assume Calzetti dust with fixed  $\epsilon = 3$ , and fit  $A_V$  with a uniform prior between 0 and 4 magnitudes. For our nebular model we fix  $a_{BC} = 10$  Myr and  $\log_{10}(U) = -3$ . In principle all of these fixed parameters could be fitted with BAGPIPES, however their marginal impact on observed photometry, particularly for quenched galaxies, and the computational expense of fitting large numbers of parameters led us to fix them to representative values, consistent with results from the literature (see Section 3.1).

We first build a baseline model of the kind commonly used in the literature (e.g. Mortlock et al. 2017). We parameterise the SFHs with a single exponentially declining component (Equation 9) and





**Figure 7.** SFH (top), mass-assembly history (middle) and normalised SFR history (bottom) for the same example object from MUFASA as shown in Fig. 4. The input histories from MUFASA are shown in blue with the posteriors shown in orange (left) and purple (right) for the exponentially declining and double-power-law SFH models of Section 4.4 respectively. The shaded regions show the 16<sup>th</sup> – 84<sup>th</sup> percentiles of the posterior. The true redshifts of quenching and of formation (as defined in Section 4.2) are overplotted as green and blue solid vertical lines respectively, with corresponding posterior median fitted values shown as dashed lines. This object was chosen as an example as it is broadly representative of the median trends visible in Fig. 8 and Fig. 9.

use uniform priors on all of our SFH parameters. We set the limits of our priors to zero and 50 Myr less than  $t(z_{\text{obs}})$  for  $T_0$ , between 300 Myr and 10 Gyr for  $\tau$  (following Wuyts et al. 2011 in both cases), between 1 and  $10^{13} M_{\odot}$  in stellar mass formed, and between 0.2 and  $2.5 Z_{\odot}$  in metallicity (Solar metallicity is assumed to take the value  $Z_{\odot} = 0.02$ ).

Once posterior distributions had been obtained for our model parameters, we generated posterior SFHs and used these to calculate the posterior distributions of  $M_*$ ,  $z_{\text{quench}}$  and  $z_{\text{form}}$  for each object. A representative comparison of MUFASA input and BAGPIPES output from fitting our exponentially declining model to mock photometry is shown in the left-hand panels of Fig. 7. At the top the SFH is shown, in the middle the mass-assembly history, and at the bottom the history of the normalised SFR (see Equation 12). It can be seen that, for this object, this model over-predicts  $M_{\text{formed}}$ , and hence  $M_*$ , and over-predicts both  $z_{\text{quench}}$  and  $z_{\text{form}}$ . There are several obvious discrepancies in the histories when viewed as a whole, with the overall shape of the SFH being poorly reproduced by the model, the mass assembly occurring much earlier than in MUFASA.

Fig. 8 shows a comparison between input values of  $M_*$ ,  $z_{\text{quench}}$  and  $z_{\text{form}}$  from MUFASA and the posterior median values obtained by fitting our exponentially declining model to the whole of the MUFASA quenched sample. It can be seen that biases are present in the estimates of our SFH parameters. This is most notable for the estimated stellar masses, which are over-predicted on average by 0.06 dex ( $\sim 15$  per cent), with the posterior median stellar masses

of  $\sim 80$  per cent of objects being overestimates. Whilst this is a relatively small bias compared to other uncertainties which exist in the determination of stellar masses (e.g. Mobasher et al. 2015), it is nevertheless significant, and interesting given the prevalence of this method for estimating stellar masses in the literature.

Both  $t(z_{\text{quench}})$  and  $t(z_{\text{form}})$  are underestimated on average by  $\approx 0.4$  Gyr. This is perhaps surprisingly good agreement, given that exponentially declining SFHs were not designed to reproduce these properties. However, as can be seen in the middle-left panel of Fig. 8, there is considerable systematic variation in the degree of underestimation of  $t(z_{\text{quench}})$  with the time between  $t(z_{\text{obs}})$  and  $t(z_{\text{quench}})$ , most notably at  $z_{\text{obs}} = 0.5$ , with the quenching times of the earliest-quenching objects actually being, on average, overestimated. This is a problem as it results in inconsistencies between quenching properties derived for different sub-samples of objects.

Motivated by these observations, we tested other SFH parameterisations and combinations of priors in order to attempt to obtain better agreement with the true values of  $M_*$ ,  $z_{\text{quench}}$  and  $z_{\text{form}}$  from MUFASA. We first investigated the effects of fitting the exponentially declining SFH parameterisation with different combinations of priors, but found no other combination that allowed us to obtain significantly better agreement with MUFASA.

We then considered the double-power-law form of Equation 10. To incorporate our prior expectation, based on the hierarchical growth of galaxies observed in cosmological simulations, that most SFHs are extended almost all the way back to the beginning of

**Table 3.** Fixed and fitted parameters with their associated priors for the double-power-law SFH model used to fit the UltraVISTA catalogue. Priors listed as logarithmic are uniform in log-space. Observed redshift values were taken from the Mortlock et al. (2017) catalogue (see Section 5). Priors placed on parameters which are well constrained by the data (e.g. stellar mass) have little to no effect on the results obtained.

Free Parameter	Prior	limits	Fixed Parameter	Value
$A_V$	Uniform	(0, 4)	$a_{BC} / \text{Myr}$	10
$\log_{10}(M_{\text{formed}} / M_{\odot})$	Uniform	(1, 13)	$\log_{10}(U)$	-3
$Z / Z_{\odot}$	Uniform	(0.2, 2.5)	$\epsilon$	3
$\tau / \text{Gyr}$	Uniform	(0, $t(z_{\text{obs}})$ )	$z_{\text{obs}}$	$z_{\text{M17}}$
$\alpha$	Logarithmic	( $10^{-2}$ , $10^3$ )	SPS models	BC03
$\beta$	Logarithmic	( $10^{-2}$ , $10^3$ )	IMF	Kroupa & Boily (2002)

cosmic time, we impose a logarithmic prior on the rising slope,  $\beta$ . In the absence of constraining data, this produces an extended, gradual rise in the SFH up to the point at which the imprint of more recent star formation becomes obvious in the shape of the observed SED. We also put a logarithmic prior on the falling slope,  $\alpha$ , however this can be easily overcome by the strong imprint which recent, rapidly quenched star formation leaves on the observed SED.

A list of the parameters and priors for our best double-power-law model are presented in Table 3. Fig. 9 shows the same comparison between input and fitted parameters for our double-power-law model as was shown for the exponentially declining model in Fig. 8. It can be seen that considerably better agreement is achieved for all three of our SFH parameters, with stellar masses now offset by only 0.02 dex. The most significant improvement is in the estimation of  $t(z_{\text{quench}})$ , where the bias on the median has been cut to 100 Myr. More significantly, the contours on the middle-left panel are now well aligned with the diagonal, indicating no change to the bias in our estimates of  $z_{\text{quench}}$  with increasing time interval between  $t(z_{\text{obs}})$  and  $t(z_{\text{quench}})$ . Additionally, the observed scatter in our estimates of  $t(z_{\text{quench}})$  have been cut by  $\approx 20$  per cent.

Fig. 4 shows example spectral and corner plots for a fit to one of the MUFASA objects using this model. This is the same object as shown in Fig. 7. Our fitted SFH, mass assembly history and normalised SFR history for our double-power-law model are shown on the right-hand panels of Fig. 7. These can be seen to be in considerably better agreement with the input histories from MUFASA than for the exponentially declining model, with close agreement being achieved with the input histories at all times.

The better agreement achieved with input MUFASA SFH parameters using our double-power-law model led us to take this model forward for use in fitting our UltraVISTA sample in Section 5. Under the assumption that the SFHs of real massive quenched galaxies look similar to those from MUFASA, we have shown that this model will allow us to obtain negligibly biased estimates of  $M_*$ ,  $z_{\text{quench}}$  and  $z_{\text{form}}$  for the galaxies in our UltraVISTA sample.

## 5 ULTRAVISTA FITTING AND SAMPLE SELECTION

In this section we describe the process of selecting our sample of 9289 massive quenched galaxies from the UltraVISTA catalogue introduced in Section 2. We begin by fitting our double-power-law model within BAGPIPES to every object in the catalogue, except for those which were identified as stars or potential AGN by Mortlock et al. (2017). Parameters and priors for our model are listed in Table 3. Fitting our BAGPIPES model whilst allowing observed redshift to vary produces more scatter in the recovered photometric redshifts than Mortlock et al. (2017) obtained by taking the median result of five photometric-redshift codes, therefore it was decided to fix our observed redshifts to these median values.

We then select objects with posterior median stellar masses,  $M_* > 10^{10} M_{\odot}$  and calculate the posterior distributions of the normalised SFR (as defined in Section 4.2) for these objects. We then select our quenched sample to contain galaxies which have posterior median normalised SFR values less than 0.1. At this stage we also exclude poorly fit objects with minimum reduced chi-squared values of greater than 3. These objects are almost exclusively the result of individual datapoints affected by large systematic errors. They comprise  $\sim 1$  per cent of the sample and are not strongly clustered in observed redshift or stellar mass.

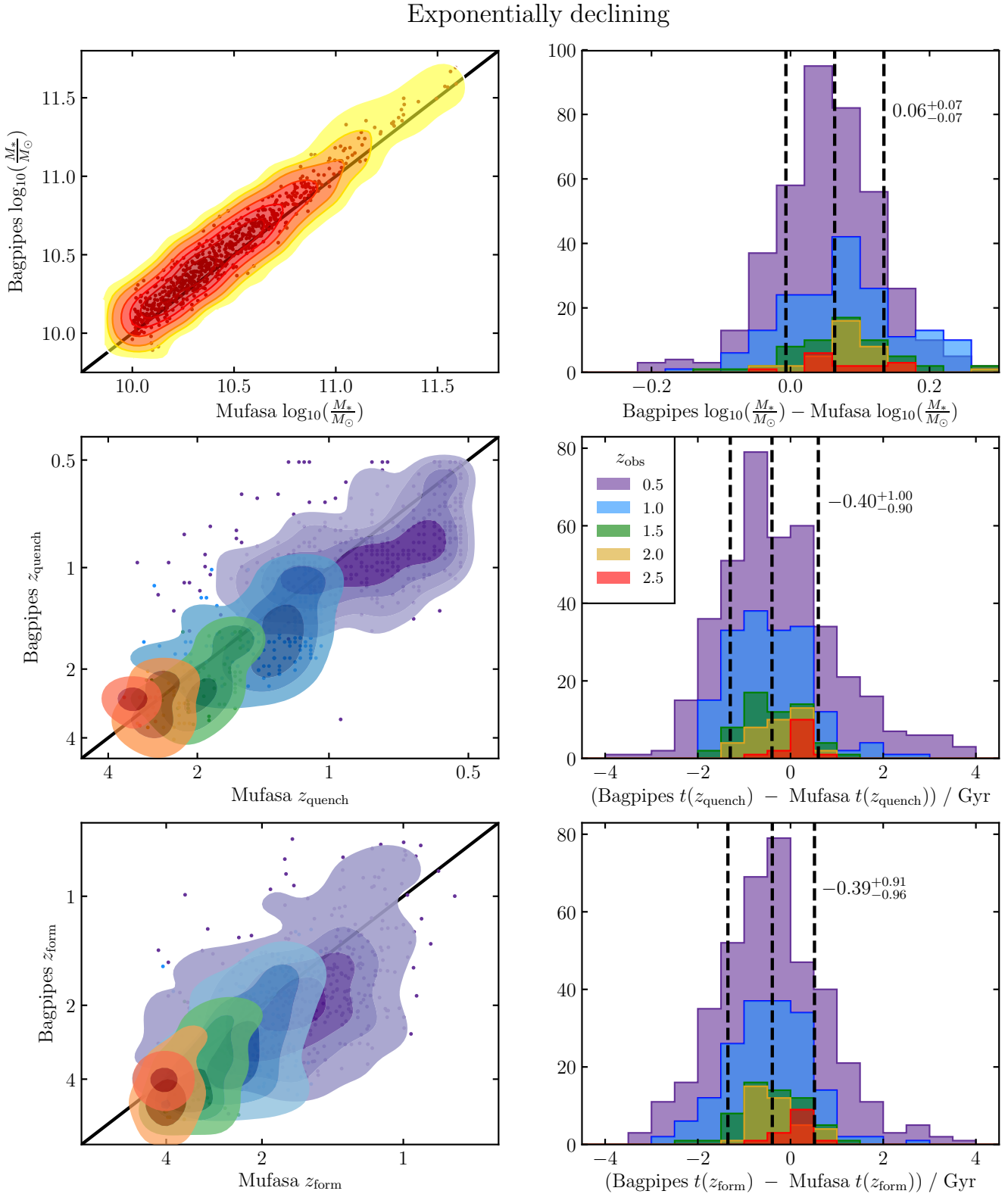
Fig. 10 shows our final sample of quenched galaxies on the UVJ plane, as well as star-forming galaxies which match all of our other selection criteria except our cut in normalised SFR. A UVJ selection of the kind proposed by Williams et al. (2009) is also plotted. Our normalised SFR selection and UVJ selection can be seen to select very similar samples of objects at all redshifts, with typically  $\gtrsim 90$  per cent agreement between the two methods. Our method can be seen to be more subtle in removing star-forming galaxies close to the UVJ boundary.

We split our sample into six bins in observed redshift between  $z = 0.25 - 3.75$ , chosen to span similar intervals in cosmic time. We then split each of these into five bins in stellar mass. Our first four mass bins have widths of 0.3 dex from  $\log_{10}(M_*/M_{\odot}) = 10 - 11.2$ , whilst our highest mass bin includes all objects with  $\log_{10}(M_*/M_{\odot}) > 11.2$ . Fig. 11 shows the distribution of our final quenched sample across our bins in stellar mass and observed redshift. Bin edges are marked in black.

The Mortlock et al. (2017) catalogue is mass complete down to  $\log_{10}(M_*/M_{\odot}) = 10$  over the UltraVISTA wide region at  $z \lesssim 1.5$ , and over the deep region at  $z \lesssim 2$ . We therefore exclude objects from our highest-redshift, lowest-mass bins if they fall below the mass-completeness limits of the UltraVISTA region they are drawn from. Bins are shaded orange in Fig. 11 for objects in the wide region having been excluded, or red for all objects having been excluded. These final criteria leave us with a sample of 9289 objects.

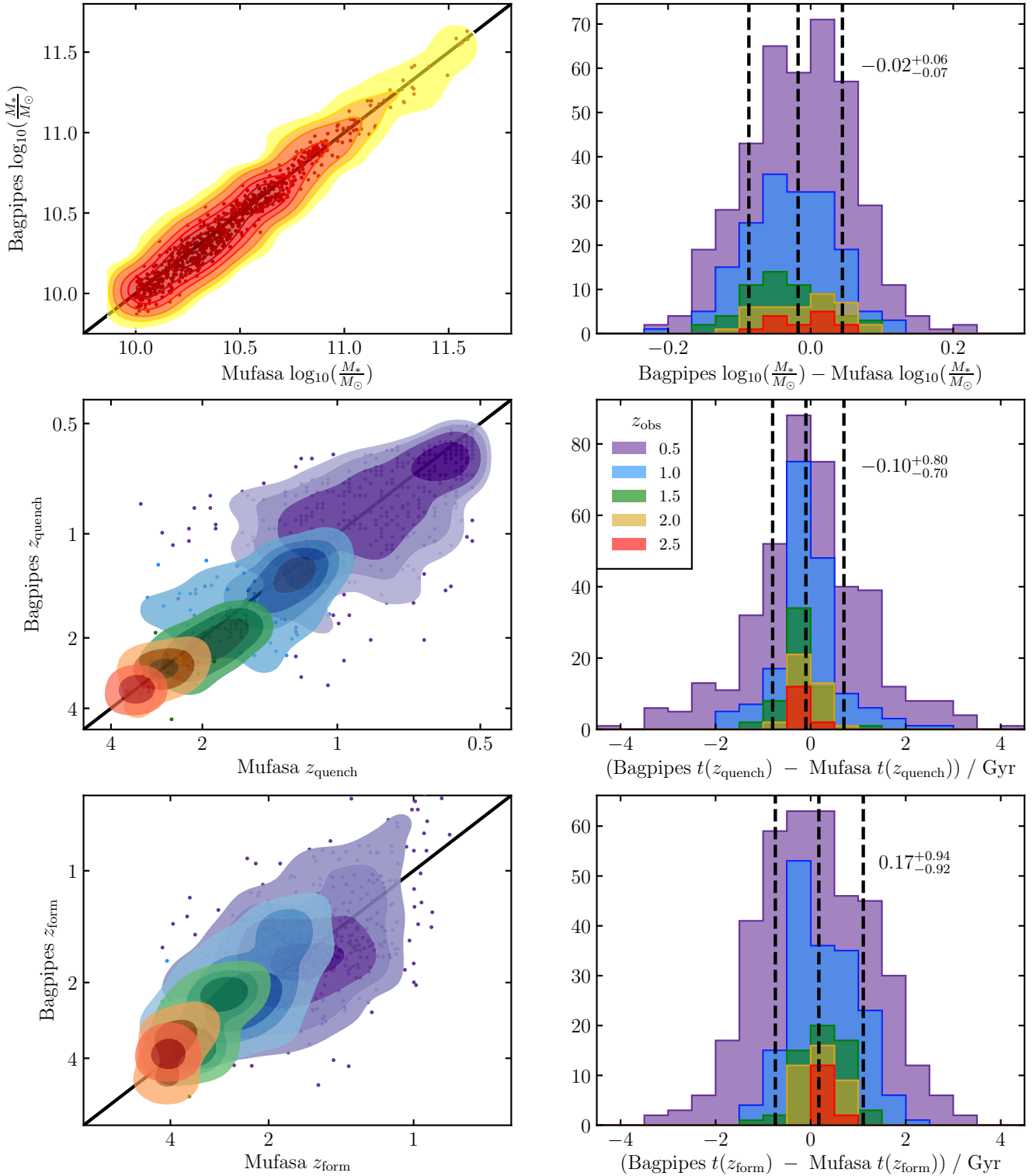
## 6 THE STAR-FORMATION HISTORIES OF ULTRAVISTA QUENCHED GALAXIES

In this section we present and discuss the results of our analysis of the SFHs of massive quenched galaxies from UltraVISTA. We analyse the SFHs using the scheme presented in Section 4.2, by calculating posterior SFHs for each of our objects and using these to calculate posterior distributions for  $t(z_{\text{form}})$ ,  $t(z_{\text{quench}})$ ,  $\Delta t_{\text{quench}}$  and  $\tau_{\text{quench}}$ . The goal of our analysis is to address Questions (i) and (ii) posed in Section 1. Hence, in Section 6.1 we consider Question (i): when did quenched galaxies form their stellar mass? Then, in Section 6.2 we consider Question (ii): how long did the process of quenching take?



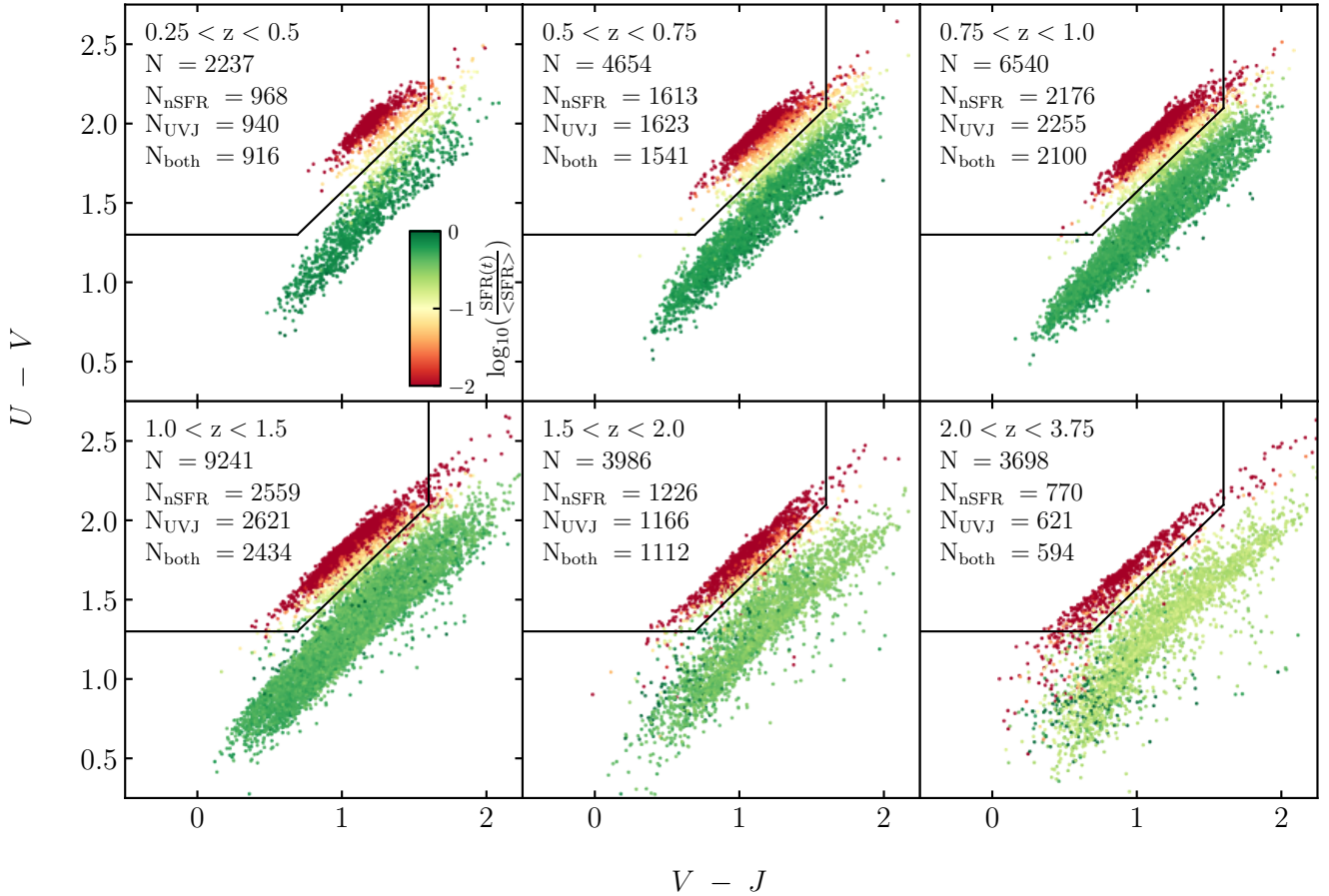
**Figure 8.** Comparison between SFH properties of MUFASA galaxies and posterior median values derived by fitting their mock photometric observations with BAGPIPES using the exponentially declining SFH parameterisation described in Section 4.4. The top-left panel shows stellar mass, middle-left panel shows redshifts of quenching and bottom-left panel shows redshifts of formation for the whole MUFASA quenched sample of 677 objects. The right-hand panels show histograms of the offset between fitted and true values of log-stellar mass (top), time of quenching (middle) and time of formation (bottom). All panels except the top-left are split by colour for different observed redshifts, as shown on the middle-right panel. The dashed lines on the right-hand panels indicate the 16th, 50th and 84th percentiles of the combined distribution for all observed redshifts. Numerical ranges for these confidence intervals are shown on the figure. The stellar masses are overestimated by  $\sim 15$  per cent on average, and both  $t(z_{\text{quench}})$  and  $t(z_{\text{form}})$  are underestimated on average by  $\approx 0.4$  Gyr.

## Double power law



**Figure 9.** Comparison between SFH properties of MUFASA galaxies and posterior median values derived by fitting their mock photometric observations with BAGPIPES using the double-power-law SFH parameterisation described in Table 3. Plot details are as in Fig. 8. Considerably better agreement can be seen for all three parameters at each observed redshift when compared to the exponentially declining SFH results displayed in Fig. 8.





**Figure 10.** UVJ selection diagram for the sample of galaxies from UltraVISTA described in Section 5. The population of star-forming galaxies which match all of our criteria except our cut in normalised SFR (nSFR; see Equation 12) are also shown. A strong trend in nSFR perpendicular to the star-forming sequence can be seen, as has been observed for sSFR (e.g. Fang et al. 2017). A UVJ selection of the kind proposed by Williams et al. (2009), with limits  $(U - V) > 1.3$ ,  $(V - J) < 1.6$ ,  $(U - V) > 0.88(V - J) + 0.69$  is overplotted in black for reference, however this was not included in our selection criteria. Our selection criterion for identifying quenched galaxies by  $\text{nSFR} < 0.1$  can be seen to be in good agreement with the UVJ selection shown in black at all observed redshifts.

As can be seen from Fig. 4, the SFHs of our individual objects are relatively poorly constrained. At this point, the question of how best to present and draw conclusions from the richness of the posterior information we have obtained naturally arises. The optimal approach to analysis of this dataset would be to construct a Bayesian hierarchical model in which the values of  $z_{\text{form}}$  and  $z_{\text{quench}}$  for each object were assumed to be drawn from some overall sample distribution with stellar-mass and observed-redshift dependencies, which could then be constrained by our whole dataset. However the computational expense of implementing this (e.g. using Gibbs sampling) was judged to be prohibitive. Instead we have elected to represent the posterior distribution for each object by the median values of relevant parameters, and to take advantage of our large sample size to make inferences about how these vary across the stellar-mass and observed-redshift bins plotted in Fig. 11.

### 6.1 When did quenched galaxies form their stellar mass?

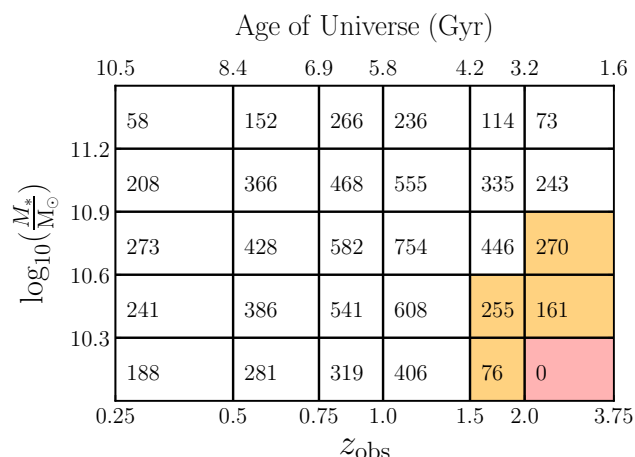
Fig. 12 shows histograms of the posterior median values of  $z_{\text{form}}$  and  $z_{\text{quench}}$  for objects in our sample, as well as the distributions of observed redshifts, split into the stellar-mass and observed-redshift bins shown in Fig. 11. The distributions of  $z_{\text{form}}$  and  $z_{\text{quench}}$  can be

seen to be, on the whole, unimodal and roughly Gaussian in shape. It is therefore reasonable to represent these distributions by their median values, which are plotted as solid vertical lines in Fig. 12. The median values of  $z_{\text{form}}$  with their standard errors are plotted against median stellar mass in Fig. 13.

Two clear trends are visible in the distributions of  $z_{\text{form}}$  from Fig 12 and Fig. 13. Firstly there is the well-known downsizing trend, in which more massive galaxies at fixed observed redshift have higher median  $z_{\text{form}}$ . This is discussed in Section 6.1.1. Secondly, there is a trend towards progressively lower median  $z_{\text{form}}$  for objects at lower observed redshifts. We discuss this in Section 6.1.2.

#### 6.1.1 Downsizing: higher formation redshifts at higher masses

The downsizing trend we observe, most clearly in Fig. 13, has also been observed in many other works (see Section 1). Siudek et al. (2017) consider a sample of quenched galaxies at  $0.4 < z < 1.0$  and show in their fig. 11 redshifts of formation calculated from stacked spectra in bins of  $\Delta z = 0.1$  across the same range of stellar masses we consider. Their redshifts of formation are considerably lower than those we show in Fig. 13 for our corresponding redshift bins. However, as they note, their methodology of assuming the SFHs



**Figure 11.** Number of galaxies in each of our bins in stellar mass and observed redshift. The edges of our bins are shown in black. Bins shaded orange are those for which UltraVISTA wide is not mass complete, and hence objects from the wide area were removed from our sample. The bin shaded red falls below the mass-completeness limit of both the wide and deep regions, and hence was excluded entirely from our sample.

of their galaxies to be composed of only a single burst component means their redshift of formation traces the time of the most recent episode of significant star formation. This means the formation redshifts they report are more analogous to our  $z_{\text{quench}}$  than our  $z_{\text{form}}$ . It can be seen from Fig. 12 that the median  $z_{\text{quench}}$  we derive for our  $0.5 < z < 0.75$  and  $0.75 < z < 1.0$  bins are consistent with the redshifts of formation shown on their fig. 11, ranging from  $z \approx 1.1$  at  $M_* \sim 10^{10} M_\odot$  to  $z \approx 1.5$  at the highest masses. It can also be seen from fig. 12 of Siudek et al. (2017) that these results are broadly consistent with analyses by Moresco et al. (2010), Jørgensen & Chiboucas (2013) and Onodera et al. (2015), whereas the results of Thomas et al. (2010) and Choi et al. (2014) confirm the continuation of the trends we observe in Fig. 13 down to  $z \approx 0$ .

Pacifici et al. (2016) perform an analysis more similar to ours than any of the studies mentioned above, both in terms of sample selection (see Fig. 5) and SFH recovery method, with the median SFHs they calculate also found to be well fitted by a double-power-law. Their redshifts of peak star formation,  $z_{\text{peak}}$  are therefore reasonably analogous to our  $z_{\text{form}}$ . In their lowest-redshift bin ( $0.2 < z < 0.5$ ) they find a range of  $z_{\text{peak}}$  between  $\approx 0.85$  at  $M_* \sim 10^{10} M_\odot$  and  $\approx 1.3$  at the highest masses; this is lower than the range of  $z \approx 1.5$  to  $1.7$  we find at similar observed redshifts in Fig. 13. This discrepancy is still apparent, though less marked at higher observed redshifts, up until  $z_{\text{obs}} \approx 1.3$ , after which our results are in good agreement. This disagreement is a consequence of differences in how the two redshifts are measured, with our  $z_{\text{form}}$  being higher than  $z_{\text{peak}}$  for SFHs which decline more rapidly than they rise (e.g. the middle panel of Fig. 6). The majority of our galaxies do follow this pattern (see Section 6.2) and so have  $z_{\text{form}}$  higher than  $z_{\text{peak}}$ .

The downsizing trend we observe is relatively weak, with only  $\approx 0.5$  Gyr in  $t(z_{\text{form}})$  separating the lowest and highest-mass galaxies in our sample at all observed redshifts. Pacifici et al. (2016) tentatively report a stronger downsizing trend at lower observed redshifts. This is not obvious in our results, with the strength of our downsizing trend across our whole observed redshift range being similar to that which they see at higher observed redshifts. However, in our lowest-redshift bin our number counts are significantly lower

than at higher redshifts, increasing the statistical uncertainties on our results, as can be seen in Fig. 13.

In summary, our UltraVISTA SFHs exhibit a downsizing trend broadly consistent with results from the literature. Our trend is not particularly strong and does not evolve significantly with observed redshift. Our median  $t(z_{\text{form}})$  evolves by  $\approx 0.5$  Gyr over a range of  $\sim 1.25$  dex in stellar mass, meaning the trend weakens as a fraction of the age of the Universe with decreasing observed redshift.

### 6.1.2 Decreasing formation redshift with decreasing observed redshift

Along with the downsizing trend discussed in Section 6.1.1, another common observation from similar analyses has been a trend towards lower average redshifts of formation with decreasing observed redshift for quiescent galaxies of all masses. This effect is observed for both  $z_{\text{form}}$  and  $z_{\text{quench}}$  in Fig. 12, and is clearly seen for  $z_{\text{form}}$  in Fig. 13. The effect is strongest at low redshift, with our lowest observed-redshift bin being strongly offset from the others, and persists across our whole observed-redshift range.

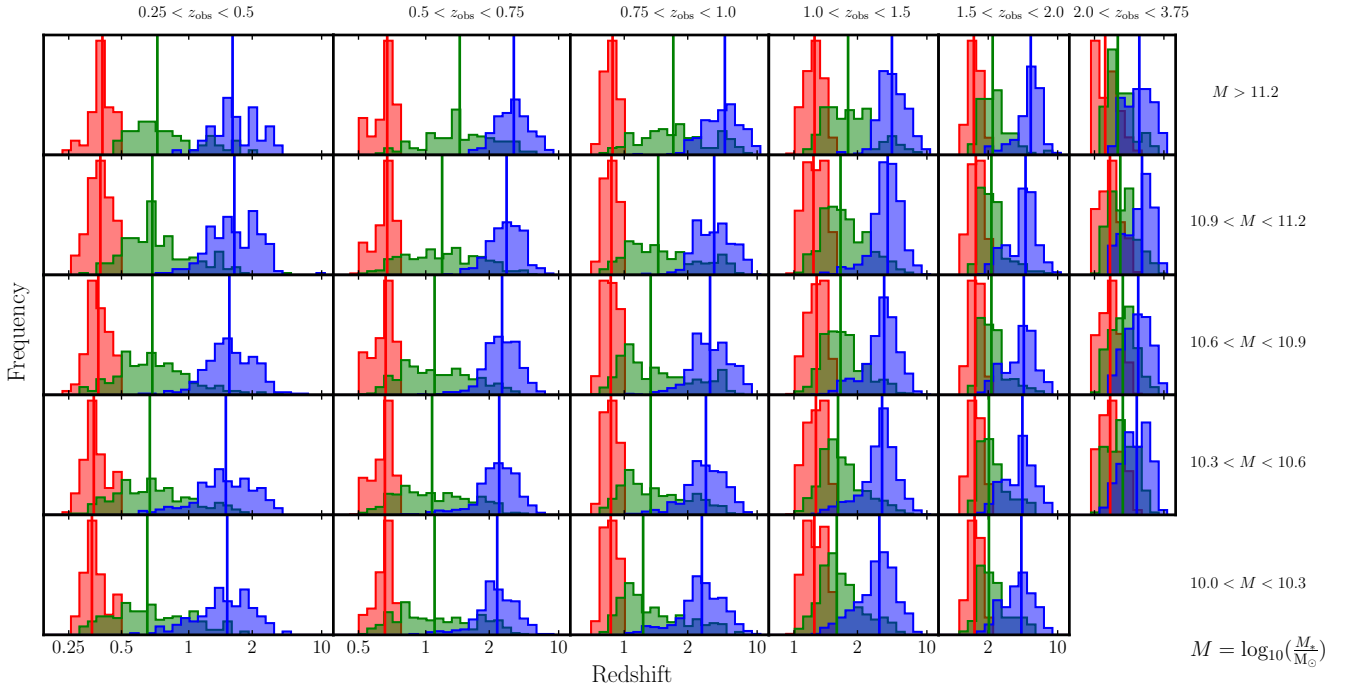
The simplest interpretation of this observation (as suggested by both Pacifici et al. 2016 and Siudek et al. 2017) relates to the continuing assembly of the red sequence throughout our observed-redshift range (e.g. Faber et al. 2007; Tomczak et al. 2014). This means that new and younger galaxies are quenching throughout our observed-redshift range. Therefore our median  $z_{\text{form}}$  and  $z_{\text{quench}}$  values at low redshift are averages over a diverse population which quenches at a wide range of redshifts. Conversely, the quenched galaxies in our higher observed-redshift bins are a biased sub-sample of these, as our selection necessarily excludes high redshift star-forming galaxies which will have quenched by lower redshifts.

The histograms plotted in Fig. 12 support this picture, with the distributions of  $z_{\text{form}}$  and  $z_{\text{quench}}$  observed to be far more extended in our lower-observed-redshift bins. The fact that the distributions of  $z_{\text{quench}}$  are extended all the way down to the low-redshift ends of each of our observed-redshift bins is consistent with the assembly of the red sequence persisting throughout our observed-redshift range.

Another possible contributory factor to this trend of decreasing median  $z_{\text{form}}$  with decreasing observed redshift would be continuing evolution of the high-redshift quiescent population post-quenching, either through periods of rejuvenated star-formation, or merger events. This would mean that even if no galaxies joined the quiescent population across our observed redshift range, the median  $z_{\text{form}}$  we calculate at lower observed redshifts would still be lower than that which we observe at  $z_{\text{obs}} > 2$ . Whilst our analysis confirms the well-known result that new galaxies join the red sequence across our observed redshift range (we see a strong increase in the number density of quiescent galaxies), we can also gain an insight into the roles of rejuvenated star-formation episodes and mergers.

A simple way of doing this is to compare the number density of quiescent galaxies we observe at high redshift (e.g.  $z_{\text{obs}} > 1.5$ ) with the number density of extremely old galaxies we observe at  $0.25 < z_{\text{obs}} < 0.5$  (those with  $z_{\text{quench}} > 1.5$ ). If all galaxies which are quenched at  $z > 1.5$  undergo purely passive evolution from that time onwards, we would expect to observe the same number density of objects in our  $0.25 < z_{\text{obs}} < 0.5$  bin with  $z_{\text{quench}} > 1.5$  as we find for quiescent galaxies at observed redshifts above 1.5.

However, if galaxies which are quenched at  $z > 1.5$  undergo further evolution after this time, we should observe fewer extremely old objects at low redshift, as newer stellar populations have been added to objects which were already quenched at  $z > 1.5$  since that time, decreasing our measured  $z_{\text{quench}}$  for these objects.



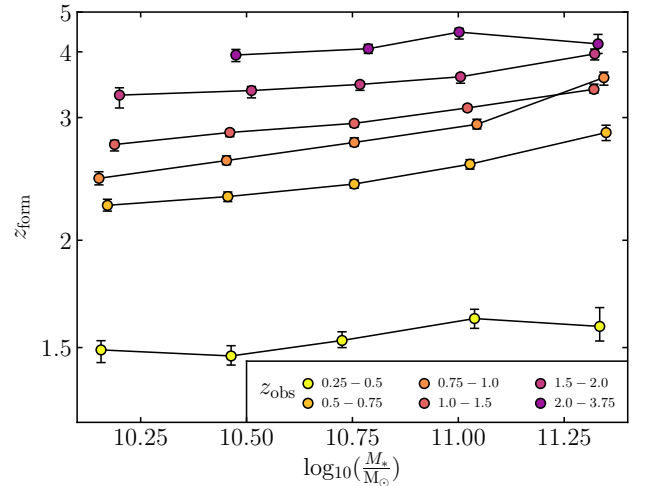
**Figure 12.** Histograms of the redshifts of observation (red), posterior median redshifts of quenching (green) and posterior median redshifts of formation (blue) for our sample of quenched galaxies, split into the stellar-mass and observed-redshift bins shown on Fig. 11. The medians of the distributions for each bin are denoted by solid vertical lines. A clear downsizing trend can be seen in both the redshifts of formation and of quenching.

We restrict this comparison to objects with  $\log_{10}(M_*/M_\odot) > 10.3$ , so that we have a mass-complete sample across our whole redshift range. Of the 780 galaxies in this mass range with  $z_{\text{obs}} < 0.5$ , just 27 have  $z_{\text{quench}} > 1.5$ , a comoving number density of  $4.7 \pm 0.9 \times 10^{-5} \text{ Mpc}^{-3}$ . By contrast we have a comoving number density of  $11.9 \pm 0.3 \times 10^{-5} \text{ Mpc}^{-3}$  for quiescent galaxies of these masses at  $z_{\text{obs}} > 1.5$ . This suggests that the majority ( $61 \pm 8$  per cent) of massive galaxies which quench at  $z_{\text{obs}} > 1.5$  undergo significant further evolution post-quenching by  $z = 0.5$ . This is in agreement with recent results from IllustrisTNG (Nelson et al. 2017), who find significant mass growth of galaxies post-quenching. It should be noted that these uncertainties take into account only Poisson noise on the number of objects observed, and not the uncertainties in  $z_{\text{quench}}$  for individual objects.

## 6.2 How long did the process of quenching take?

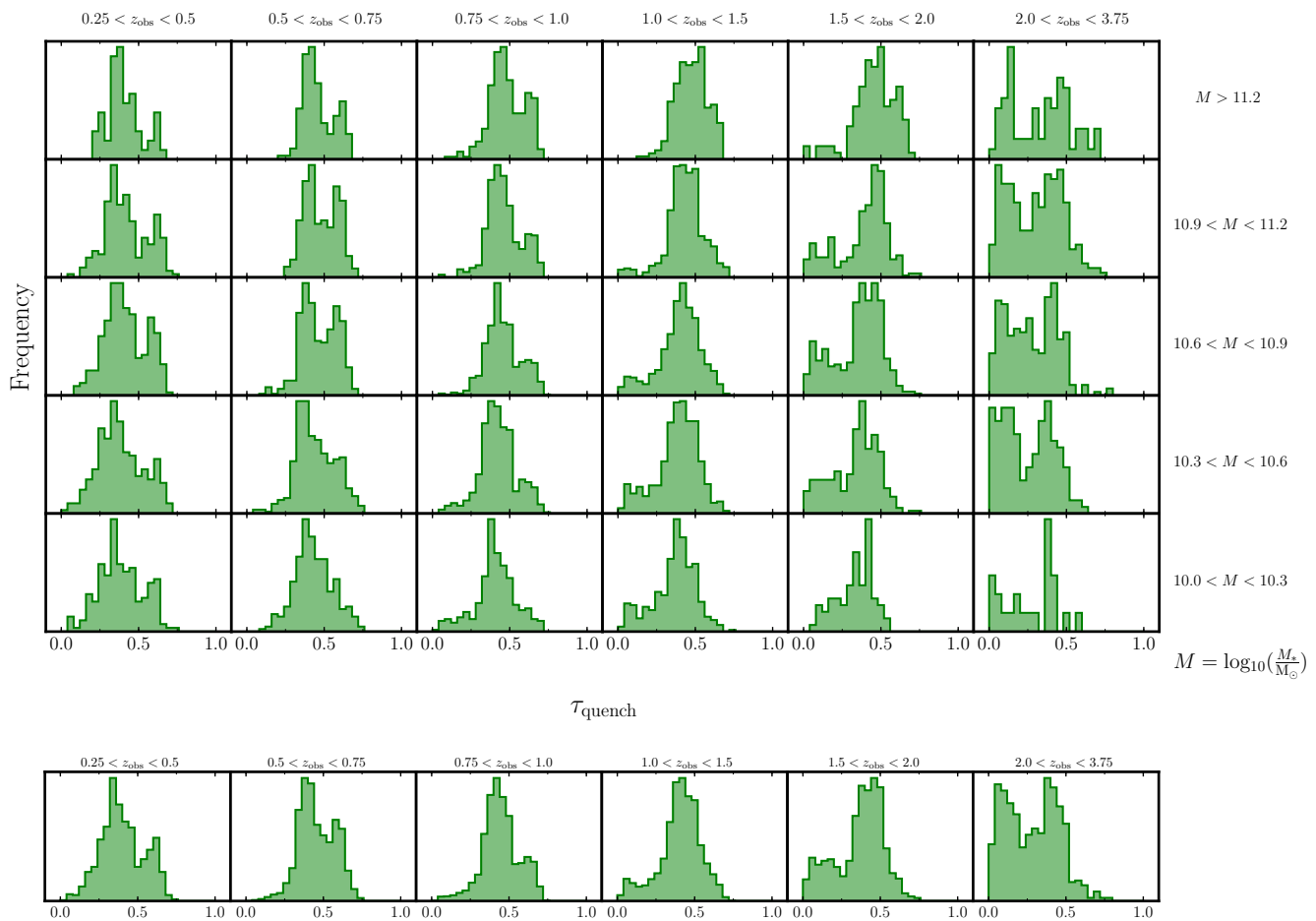
A key piece of information necessary to constrain the physical mechanisms by which galaxies quench their star formation is the timescale over which quenching takes place. Accurate recovery of SFHs is the ideal method for studying quenching timescales, however, as discussed in Section 4.1, recovering physically realistic SFHs is challenging. Even if this can be achieved there are questions as to which parameters of the SFH are most representative of the quenching timescale, usually taken to mean the timescale over which the galaxy crosses the green valley, i.e. between an initial position on the star-forming main sequence and some threshold defining passivity.

Recently, a common approach has been to fit SFH models similar to the exponentially declining model of Equation 9 to spectral indices such as D4000 and the strength of the H $\delta$  feature (e.g. Nogueira-Cavalcante et al. 2018), or UV-optical colours derived



**Figure 13.** Median  $z_{\text{form}}$  for each of our mass and redshift bins with their standard errors. Two clear trends are visible. Firstly, at fixed observed redshift, galaxies with higher stellar masses are seen to have higher median  $z_{\text{form}}$ . This is the commonly observed downsizing trend. Secondly, the median redshift of formation decreases with observed redshift at all masses.

from broad-band photometry (e.g. Smethurst et al. 2017). The timescale over which the SFR decreases,  $\tau$  is then used as a proxy for quenching timescale. However, when fitting this kind of SFH model, strong degeneracies between  $\tau$  and  $T_0$  mean that constraining  $\tau$  is extremely challenging. Also, the effects of the age-metallicity degeneracy on D4000 and H $\delta$ , and the age-metallicity-dust degeneracy



**Figure 14.** Distributions of posterior median values of  $\tau_{\text{quench}}$  (quenching timescale as a fraction of the age of the Universe when the galaxy quenched) across our bins in stellar mass and observed redshift. The bottom row shows the results from the grid collapsed along the stellar-mass axis for improved statistics.

acy on UV-optical colours must be properly treated, or the derived uncertainties in  $\tau$  will be significantly underestimated.

In Section 4.2 we define an alternative scheme, in which the quenching timescale,  $\Delta t_{\text{quench}}$  is the time interval between the formation redshift,  $z_{\text{form}}$  (corresponding to the mass-weighted age of the galaxy) and the redshift at which the SFR of the galaxy falls below 10 per cent of its average value across the SFH of the galaxy,  $z_{\text{quench}}$ . We finally define the normalised quenching timescale,  $\tau_{\text{quench}}$  to be  $\Delta t_{\text{quench}}$  as a fraction of the age of the Universe at  $z_{\text{quench}}$ . This is the fraction of the age of the Universe at the time of quenching which the galaxy takes to quench.

In Section 4.4, we confirm the ability of our double-power-law BAGPIPES model to accurately recover both  $z_{\text{form}}$  and  $z_{\text{quench}}$ . Hence we can be confident that, on average, we recover the correct values of  $\tau_{\text{quench}}$  for our UltraVISTA objects. This allows us to obtain information about the distribution of shapes their SFHs take and how these vary across the stellar-mass and observed-redshift bins shown in Fig. 11.

Fig. 14 shows the distributions of posterior median values of  $\tau_{\text{quench}}$  for objects in each of our stellar-mass and observed-redshift bins. Additionally, stacked histograms across all stellar-mass bins within a given observed-redshift bin are shown below the main grid. In Section 6.2.1 we discuss trends in these histograms across our stellar-mass and observed-redshift bins, and what these reveal about trends in the shapes of the SFHs. In Section 6.2.2 we perform a check

to understand the effects of our priors on these trends. In Section 6.2.3 we consider what these results reveal about the quenching mechanisms that have influenced the SFHs of the galaxies in our quenched sample.

### 6.2.1 Trends in SFH shape revealed by $\tau_{\text{quench}}$

Firstly, in our highest observed-redshift bin,  $2.0 < z_{\text{obs}} < 3.75$ , there is a strong peak in the distribution of  $\tau_{\text{quench}}$  centred on  $\tau_{\text{quench}} \approx 0.1$  in each of our mass bins. This peak is also apparent, though reduced in strength, across our two redshift bins between  $1.0 < z_{\text{obs}} < 2.0$ , and begins to take on a more obvious mass dependency, being observed most strongly at lower masses. A tail of small numbers of objects with  $\tau_{\text{quench}} < 0.2$  is apparent down to  $z = 0.25$  at low masses. In contrast, virtually no objects are observed to fall into this region above  $M_* = 10^{10.5} M_{\odot}$  and below  $z_{\text{obs}} = 1$ . A  $\tau_{\text{quench}}$  of 0.1 corresponds to a quenching timescale of  $\sim 300$  Myr at  $z_{\text{quench}} = 2$ , or  $\sim 600$  Myr at  $z_{\text{quench}} = 1$ .

As demonstrated by the top panel of Figure 6, in order to reach such small values of  $\tau_{\text{quench}}$ , SFHs must not only be rapidly quenched, they must also rise rapidly at late times, just before they quench. This shows that at  $z_{\text{obs}} \gtrsim 1$  there is a significant component of the quenched population with SFHs that both rise and fall rapidly, in a small fraction of the age of the Universe at their time



of quenching. There is also a small component with this SFH shape at lower redshifts ( $z \lesssim 1$ ) in our lower-mass bins.

Secondly, across all of our bins in stellar mass and observed redshift there is a strong peak centred around  $\tau_{\text{quench}} \approx 0.4$ . This accounts for the bulk of objects in our sample, and indicates an extended period over which star formation rises, followed by a relatively abrupt shutoff (e.g. the middle panel of Fig. 6). These SFHs have quenching timescales of just under half the age of the Universe at their time of quenching (e.g.  $\sim 1$  Gyr at  $z_{\text{quench}} = 2$ , or  $\sim 2$  Gyr at  $z_{\text{quench}} = 1$ ), almost as rapid as the previous category of objects. However they are distinct from the previous category as they undergo a much more extended period of rising star formation before they reach their peak.

Finally, a third peak begins to emerge at  $z_{\text{obs}} < 1.0$ , centred around  $\tau_{\text{quench}} \approx 0.6$ . This population is more apparent at  $0.5 < z_{\text{obs}} < 1.0$  in our higher-mass bins, but by our lowest-redshift bin this peak is clearly visible even in our lowest-mass bin. These objects differ significantly from the previous two populations in that their SFHs rise more quickly than they fall (e.g. the bottom panel of Fig. 6), resembling the exponentially declining SFHs traditionally used in SED fitting, or the log-normal SFHs of Gladders et al. (2013), but represent a minority of our observed sample. These objects have quenching timescales greater than half of the age of the Universe at  $z_{\text{quench}}$  (e.g.  $\sim 3$  Gyr at  $z_{\text{quench}} = 1$ ).

These trends are more clearly visible in the bottom row of Fig. 14, where mass resolution has been sacrificed in order to obtain better statistics. The strongest peak at  $\tau_{\text{quench}} \sim 0.4$  is clearly dominant, whilst the presence and gradual decline of the population below  $\tau_{\text{quench}} = 0.2$  is clear at  $z_{\text{obs}} > 1$ , and the presence and gradual rise of the peak centred on  $\tau_{\text{quench}} = 0.6$  is clear at  $z_{\text{obs}} < 1$ .

Initially, Figure 14 appears to show surprisingly little variation in quenching timescale with stellar mass. However, it is important to remember that  $\tau_{\text{quench}}$  is normalised by the age of the Universe at  $z_{\text{quench}}$ . As can be seen from Figure 12, more massive galaxies have higher median  $z_{\text{quench}}$  than less massive ones. This means that Figure 14 is actually showing more rapid quenching for more massive galaxies.

### 6.2.2 A check on the effects of our priors

Considerable effort has already been made in Section 4 to understand the dependence of these results on the model we have used to fit our observed photometry. Here, we perform one further test to directly check for potential biasing effects of the priors of our SFH parameterisation on  $\tau_{\text{quench}}$ . To do this we draw SFH parameters at random from the prior distributions listed in Table 3 and calculate  $t(z_{\text{form}})$ ,  $t(z_{\text{quench}})$  and  $\tau_{\text{quench}}$  for a statistical sample of SFHs.

The value of  $\tau_{\text{quench}}$  depends only on the shape of the SFH, not its normalisation, therefore the prior distribution on  $\tau_{\text{quench}}$  is independent of stellar mass. On the top row of Fig. 15, our prior distributions are plotted in purple. Only a very weak dependence on observed redshift is seen, meaning the trends we observe across our stellar-mass and observed-redshift bins cannot be a result of biasing effects of our prior distribution.

It is still important however to understand how the overall shape of our distribution is affected by our choice of priors. As can be seen from Fig. 15, our prior distribution has strong, narrow peaks at  $\tau_{\text{quench}} = 0$  and  $0.5$ . Upon experimentation it was found that these are due to our allowing  $\alpha$  and  $\beta$  for our double-power-law SFHs (see Equation 10) to reach very low and very high values respectively. Whilst these peaks in the prior do not correspond with any of the

peaks discussed in Section 6.2.1, we still wished to understand how they might affect our results.

In order to check that these peaks do not strongly affect our results, we reduced the range of allowed  $\alpha$  and  $\beta$  values from (0.01, 1000) to (0.1, 100), which was found through experimentation to produce a relatively flat prior on  $\tau_{\text{quench}}$ . We then re-fitted our sample using this reduced range of allowed  $\alpha$  and  $\beta$  values. This new prior distribution is shown in purple on the bottom row of Fig. 15, along with the results which were obtained by fitting with this new prior. It can be seen that the structure described in Section 6.2.1 is still clearly visible under this new, flattened prior.

### 6.2.3 Physical interpretation of trends in SFH shape

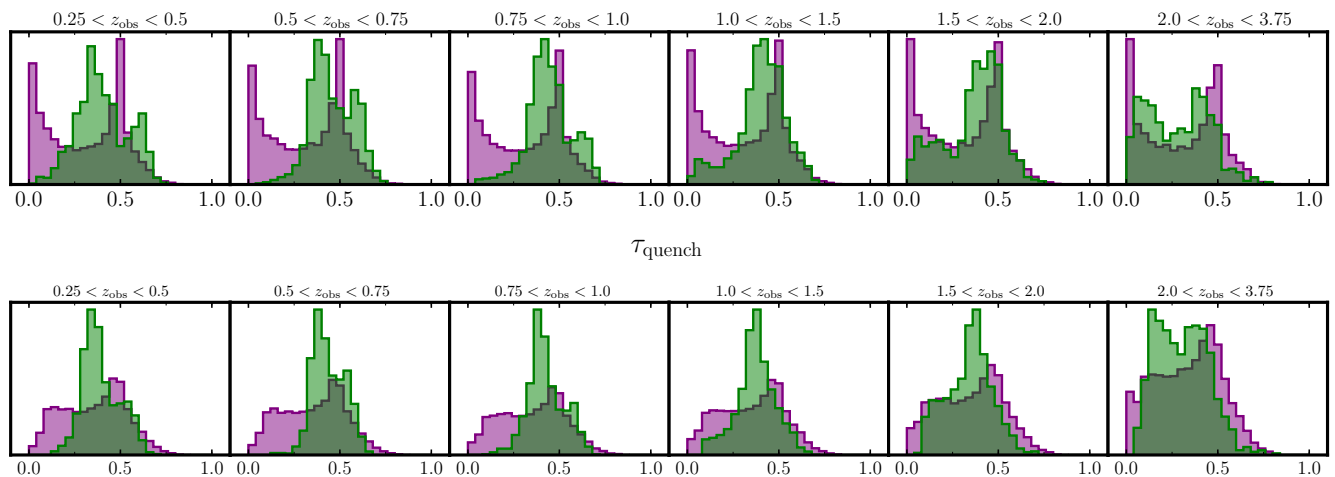
To briefly recap Section 6.2.1, we observe three peaks in the distributions of  $\tau_{\text{quench}}$  in Fig. 14, corresponding to three distinct SFH shapes. We refer to these three cases as (a), (b) and (c), which correspond roughly to the top, middle and bottom SFHs shown in Fig. 6 respectively. To summarise their properties, we have:

- (a) A peak centred on  $\tau_{\text{quench}} \approx 0.1$ , corresponding to SFHs which rise and decline rapidly, in a small fraction of the age of the Universe at  $z_{\text{quench}}$ . Strongly visible at  $z_{\text{obs}} > 1$ , also present in lower-mass bins at lower redshift. Physical quenching timescales of  $< 1$  Gyr.
- (b) A peak centred on  $\tau_{\text{quench}} \approx 0.4$ , corresponding to SFHs which decline more rapidly than they rise. Strongly visible in all of our bins of stellar mass and observed redshift, making up the majority of our observed sample. Physical quenching timescales of  $1 - 2$  Gyr.
- (c) A peak centred on  $\tau_{\text{quench}} \approx 0.6$  corresponding to SFHs which rise more rapidly than they decline. Only visible at  $z_{\text{obs}} < 1$  and most visible at higher masses, making up a further minority of our observed sample. Physical quenching timescales of  $\gtrsim 3$  Gyr.

As these SFH types are seen to be clearly distinct from each other in Fig. 14, it is reasonable to try to associate them with qualitatively different quenching processes. This is possible through comparing the physical quenching timescales we infer for each of these cases to predictions for the behaviour of different physical processes.

Case (a) is associated with a process which rapidly quenches star formation, on timescales of  $< 1$  Gyr. This process is most active at  $z_{\text{obs}} > 1$ , and is also associated with a rapid rise in star-formation activity immediately preceding quenching. However, these SFHs need not necessarily be the result of a monolithic-collapse scenario. As discussed in Section 4.4, more recent star formation leaves a more significant impact on the observed spectrum. Therefore SFHs with a slow, shallow rise from the beginning of cosmic time, followed by a large, rapidly rising and quenching burst will be fitted with something like case (a) if a double-power-law form is used. This is because the more recent star-formation episode dominates the observed spectrum. More strongly constraining data, such as rest-frame ultraviolet spectroscopy, would be needed to capture the detail of the early-time evolution of such SFHs.

The speed of quenching in case (a) is indicative of an ejective quenching process (see Section 1), as shutting off the supply of new gas would not be sufficient to cause the almost immediate cessation of star-formation activity. Quasar-mode feedback is one process which has the capability to cause extremely rapid quenching, and, if triggered by major-merger events, would also be expected to be much more common at high redshift, as the rate of major mergers is known to be a strong function of redshift (e.g. Lotz et al. 2011).



**Figure 15.** Distributions of  $\tau_{\text{quench}}$  in bins of observed redshift. The top row is the same as shown at the bottom of Fig. 14 with the prior distribution shown in purple. The bottom row shows the prior distribution and posterior median  $\tau_{\text{quench}}$  values for the case in which the prior limits on  $\alpha$  and  $\beta$  have been reduced to (0.1, 100) from their original ranges of (0.01, 1000). This change, whilst flattening the prior distribution of  $\tau_{\text{quench}}$ , does not significantly affect our results.

The merger-triggered, quasar-mode feedback scenario would also explain the rapid rise of star-formation activity which is observed immediately before these galaxies quench.

Virtually no examples of this kind of SFH are observed in our lowest-redshift bins, except at the lowest masses, supporting the conclusion we reach in Section 6.1.2 that the majority of galaxies which are observed to be quenched at high redshift undergo further interactions which affect the shapes of their SFHs when they are observed later in cosmic time. This is consistent with the findings of Gabor et al. (2011), who conclude that systems which quench due to AGN-driven outflows should begin to re-accrete gas and resume star formation within 1–2 Gyr.

The tidal/ram-pressure stripping experienced by galaxies falling into clusters is another process which can generate SFHs with this shape (e.g. Diemer et al. 2017). This could be responsible for the tail of objects at low masses at  $z_{\text{obs}} < 1$  with this kind of SFH, as the most massive galaxies are unlikely to become satellites, and mergers are less common later in cosmic time.

Case (b) is the dominant mode of quenching we observe within our sample, and is associated with a process which quenches star formation on intermediate timescales of  $\sim 1$ –2 Gyr. The dominant mode of quenching in modern hydrodynamic simulations such as MUFASA (Davé et al. 2016) and IllustrisTNG (Nelson et al. 2017) is AGN-driven feedback in the low-accretion (jet) mode, the implementation of which has been shown to significantly improve agreement with the observed properties of the  $z = 0$  red sequence (e.g. Gabor & Davé 2015). The typical timescale for quenching due to this process in IllustrisTNG is  $\sim 1.6$  Gyr, which is consistent with what we observe for case (b). We therefore tentatively associate case (b) with AGN-driven feedback in the low-accretion mode.

Finally, case (c) is associated with some process which leads to a gradual decline in star formation over a period of  $\gtrsim 3$  Gyr. Clearly this process must be preventative, rather than ejective, however it is the most challenging to associate with a known physical mechanism. As previously noted, this kind of SFH is similar in shape to an exponentially declining or log-normal function. This, and the fact that these objects are observed earliest ( $z \approx 1$ ) at the highest stellar masses, suggests a scenario similar to the model of Gladders et al. (2013). In this model, star formation in individual galaxies naturally

dies down in the same way as the cosmic SFR density as the cosmic gas supply gradually dwindles. In the Gladders et al. (2013) model, the rate is determined, in part, by stellar mass, with the most massive galaxies reaching the end-point of their evolution the fastest.

Having tentatively identified the quenching processes at work, it is interesting to consider how their relative importances vary as a function of stellar mass. As was noted in Section 6.2.1, the distributions of  $\tau_{\text{quench}}$  are remarkably consistent across our bins in stellar mass, which translates into faster physical quenching timescales for more massive galaxies, as we find these to be quenched at higher redshifts on average (Figure 11).

This finding is in agreement with previous work (e.g. Thomas et al. 2010; Pacifici et al. 2016). However the consistency of our distributions of  $\tau_{\text{quench}}$  across our stellar mass bins suggests that faster quenching at higher stellar masses is a further consequence of mass-accelerated evolution (or downsizing), rather than a consequence of a change in quenching mechanism across the range of masses we probe ( $M_* > 10^{10} M_{\odot}$ ). The fact that we first begin to see case (c) galaxies (which seem to be associated with a more advanced stage in galaxy evolution) first at the highest masses lends further weight to this theory. The same argument applies to the fact that we observe case (a) galaxies down to lower observed redshifts only at lower masses. These less-massive galaxies appear to be following their more-massive counterparts along the same evolutionary track at a slower pace.

## 7 CONCLUSIONS

We present the new galaxy spectral model generation framework and fitting tool Bayesian Analysis of Galaxies for Physical Inference and Parameter Estimation, or BAGPIPES, which can be used to rapidly generate complex, physically realistic model galaxy spectra and fit these to arbitrary combinations of spectroscopic and photometric data using the MULTINEST nested sampling algorithm.

We verify our ability to recover the properties of the SFHs of quenched galaxies by fitting mock observations for a sample of simulated massive quenched galaxies with realistic SFHs, chemical-enrichment histories and dust properties from the MUFASA suite of

cosmological hydrodynamic simulations. We investigate thoroughly the effects of our parameterisation and priors on the biases in our derived SFH properties. We find that the use of an exponentially declining SFH model produces a small but significant overestimation of the stellar masses of quenched galaxies, and an overestimation of the redshifts at which their stellar masses were assembled. A significant improvement is observed when fitting a double-power-law SFH model with the priors listed in Table 3.

We then use BAGPIPES to perform a detailed analysis of the SFHs of a large sample of 9289 quenched galaxies from the UltraVISTA Survey with stellar masses,  $M_* > 10^{10} M_\odot$  across the observed redshift range  $0.25 < z_{\text{obs}} < 3.75$ . We make the following observations based on our analysis:

(i) We observe a clear downsizing trend in our results (Fig. 13), with the most massive galaxies being on average  $\approx 0.5$  Gyr older than those with  $M_* \approx 10^{10} M_\odot$ . This appears to remain relatively constant across our whole observed-redshift range.

(ii) A simple analysis of the redshift evolution of the number density of galaxies with very old stellar populations (Section 6.1.2) suggests that the majority ( $61 \pm 8$  per cent) of galaxies which quench at  $z > 1.5$  will undergo significant further evolution through rejuvenated star-formation or merger events by  $z = 0.5$ .

(iii) The SFHs of our quenched galaxies fall into three distinct classes with different shapes (Fig. 14). We refer to these as cases (a), (b) and (c), which roughly correspond to the top, middle and bottom panels of Fig. 6 respectively. We propose that these different cases correspond to different quenching mechanisms (Section 6.2.3).

(iv) Galaxies with SFHs of type (a) are common at  $z_{\text{obs}} > 1$ . They experience a rapid rise followed by a rapid decline in SFR over a small fraction of the age of the Universe at the redshift of their quenching, and have rapid quenching timescales of  $< 1$  Gyr. We tentatively identify these as the result of (possibly merger-triggered) quasar-mode AGN feedback.

(v) Galaxies with SFHs of type (b) make up the bulk of our quenched sample across all redshifts and masses. They experience a slower rise and more rapid decline in SFR, with intermediate quenching timescales of  $1 - 2$  Gyr. These SFHs match predictions from simulations for quenching by low-accretion (jet) mode AGN feedback (e.g. Davé et al. 2017; Nelson et al. 2017).

(vi) Galaxies with SFHs of type (c) begin to enter our sample at  $z_{\text{obs}} < 1$ . Their SFHs rise more rapidly than they fall, and have long quenching timescales of  $\gtrsim 3$  Gyr. These are not obviously associated with a specific physical process, and we speculate that they may be the first galaxies in which star formation ‘naturally’ dies down with the decrease of the overall cosmic gas supply.

(vii) In common with other studies, we observe more rapid quenching at higher stellar masses. However, in contrast to some previous studies, we see no clear evidence of a change in the relative importances of different quenching mechanisms with stellar mass (at masses greater than  $M_* = 10^{10}$ ). Our results suggest that this trend is an extension of the well-known mass-accelerated evolution (or downsizing) trend, which appears to affect quenching timescale as well as the redshift of formation.

To summarise, our results suggest that the red sequence at  $M_* > 10^{10} M_\odot$  has been built up since  $z \sim 4$  by a number of different quenching processes which make different relative contributions at different epochs. At  $z \gtrsim 1$ , (potentially merger-triggered)

quasar-mode AGN feedback plays a significant role in the rapid but short-lived quenching of galaxies. Throughout cosmic time, AGN feedback in the low-accretion mode is the dominant quenching mechanism, being responsible for the bulk of the build-up of the red sequence. However galaxies which quench by this processes at  $z \gtrsim 1$  typically either experience some periods of rejuvenated star formation at later times or merge with younger stellar populations which further reduces the number of extremely early-quenching objects observed at low redshift. Finally, at  $z \lesssim 1$ , a third mode of ‘natural’ quenching begins to act, shutting down star-formation much more slowly as the cosmic gas supply declines. It is also possible that at lower masses we see rapid quenching due to tidal/ram-pressure stripping of satellite galaxies. This is qualitatively similar to the model of Schawinski et al. (2014). A joint analysis combining morphology and environment with detailed SFH recovery would provide an additional test of the scheme we propose.

In this work we have considered a large sample of galaxies with high quality photometric data. These data are nevertheless weakly constraining on the SFHs of individual objects (e.g. Fig. 4). We have relied instead upon our large sample size to obtain information about the distribution of the SFH properties of quenched galaxies across bins in stellar mass and observed redshift. This was motivated, in part, by a desire to establish a baseline for how well the SFHs of quenched galaxies could be constrained using photometric observations alone.

Recently, several authors (e.g. Pacifici et al. 2012; Thomas et al. 2017) have confirmed that significantly better constraints on physical parameters can be obtained by fitting combinations of rest-frame ultraviolet spectroscopy and multi-wavelength photometry. In an upcoming paper we will extend our analysis to fitting spectroscopy from VANDELS (McLure et al. 2018; Pentericci et al. 2018) with BAGPIPES, in order to obtain more detailed information on the properties of individual SFHs, to build up a more detailed understanding of the quenching processes which affect these galaxies.

## ACKNOWLEDGEMENTS

The authors all acknowledge the support of the UK Science and Technology Facilities Council. This work is based on data products from observations made with ESO Telescopes at Paranal Observatory under ESO programme ID 179.A2005, and on data products produced by TERAPIX and the Cambridge Astronomy survey Unit on behalf of the UltraVISTA consortium. This work is based in part on observations made with the Spitzer Space Telescope, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under a NASA contract. This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013).

## References

- Abramson L. E., Gladders M. D., Dressler A., Oemler Jr. A., Poggianti B., Vulcani B., 2016, *ApJ*, **832**, 7
- Anders E., Grevesse N., 1989, *Geochimica Cosmochimica Acta*, **53**, 197
- Ashby M. L. N., et al., 2013, *ApJ*, **769**, 80
- Astropy Collaboration et al., 2013, *A&A*, **558**, A33
- Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, *MNRAS*, **373**, 469
- Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, **93**, 5
- Barro G., et al., 2013, *ApJ*, **765**, 104
- Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, **770**, 57

- Bell E. F., et al., 2004, *ApJ*, **608**, 752
- Bell E. F., et al., 2012, *ApJ*, **753**, 167
- Brammer G. B., et al., 2009, *ApJ*, **706**, L173
- Brinchmann J., Charlot S., Kauffmann G., Heckman T., White S. D. M., Tremonti C., 2013, *MNRAS*, **432**, 2112
- Bruzual G., Charlot S., 2003, *MNRAS*, **344**, 1000
- Buchner J., et al., 2014, *A&A*, **564**, A125
- Byler N., Dalcanton J. J., Conroy C., Johnson B. D., 2017, *ApJ*, **840**, 44
- Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, **533**, 682
- Cardelli J. A., Clayton G. C., Mathis J. S., 1989, *ApJ*, **345**, 245
- Carson D. P., Nichol R. C., 2010, *MNRAS*, **408**, 213
- Casey C. M., 2012, *MNRAS*, **425**, 3094
- Charlot S., Fall S. M., 2000, *ApJ*, **539**, 718
- Charlot S., Longhetti M., 2001, *MNRAS*, **323**, 887
- Chevallard J., Charlot S., 2016, *MNRAS*, **462**, 1415
- Choi J., Conroy C., Moustakas J., Graves G. J., Holden B. P., Brodwin M., Brown M. J. I., van Dokkum P. G., 2014, *ApJ*, **792**, 95
- Chuter R. W., et al., 2011, *MNRAS*, **413**, 1678
- Cid Fernandes R., Mateus A., Sodré L., Stasińska G., Gomes J. M., 2005, *MNRAS*, **358**, 363
- Citro A., Pozzetti L., Moresco M., Cimatti A., 2016, *A&A*, **592**, A19
- Croton D. J., et al., 2006, *MNRAS*, **365**, 11
- Cullen F., et al., 2017a, preprint, ([arXiv:1712.01292](https://arxiv.org/abs/1712.01292))
- Cullen F., McLure R. J., Khochfar S., Dunlop J. S., Dalla Vecchia C., 2017b, *MNRAS*, **470**, 3006
- Davé R., Thompson R., Hopkins P. F., 2016, *MNRAS*, **462**, 3265
- Davé R., Rafieferantsoa M. H., Thompson R. J., 2017, *MNRAS*, **471**, 1671
- Diemer B., Sparre M., Abramson L. E., Torrey P., 2017, *ApJ*, **839**, 26
- Dopita M. A., Kewley L. J., Heisler C. A., Sutherland R. S., 2000, *ApJ*, **542**, 224
- Eldridge J. J., Stanway E. R., 2009, *MNRAS*, **400**, 1019
- Faber S. M., et al., 2007, *ApJ*, **665**, 265
- Falcón-Barroso J., Sánchez-Blázquez P., Vazdekis A., Ricciardelli E., Cardiel N., Cenarro A. J., Gorgas J., Peletier R. F., 2011, *A&A*, **532**, A95
- Fang J. J., et al., 2017, preprint, ([arXiv:1710.05489](https://arxiv.org/abs/1710.05489))
- Ferland G. J., et al., 2017, preprint, ([arXiv:1705.10877](https://arxiv.org/abs/1705.10877))
- Feroz F., Hobson M. P., 2008, *MNRAS*, **384**, 449
- Feroz F., Hobson M. P., Bridges M., 2009, *MNRAS*, **398**, 1601
- Feroz F., Hobson M. P., Cameron E., Pettitt A. N., 2013, preprint, ([arXiv:1306.2144](https://arxiv.org/abs/1306.2144))
- Foreman-Mackey D., 2016, *The Journal of Open Source Software*, **24**
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, **125**, 306
- Förster Schreiber N. M., et al., 2014, *ApJ*, **787**, 38
- Fumagalli M., et al., 2016, *ApJ*, **822**, 1
- Furusawa H., et al., 2016, *ApJ*, **822**, 46
- Gabor J. M., Davé R., 2015, *MNRAS*, **447**, 374
- Gabor J. M., Davé R., Finlator K., Oppenheimer B. D., 2010, *MNRAS*, **407**, 749
- Gabor J. M., Davé R., Oppenheimer B. D., Finlator K., 2011, *MNRAS*, **417**, 2676
- Gallazzi A., Bell E. F., Zibetti S., Brinchmann J., Kelson D. D., 2014, *ApJ*, **788**, 72
- Gladders M. D., Oemler A., Dressler A., Poggianti B., Vulcani B., Abramson L., 2013, *ApJ*, **770**, 64
- Goodman J., Weare J., 2010, *Communications in Applied Mathematics and Computational Science*, Vol.-5, No.-1, p.-65-80, 2010, **5**, 65
- Hartley W. G., et al., 2010, *MNRAS*, **407**, 1212
- Heavens A. F., Jimenez R., Lahav O., 2000, *MNRAS*, **317**, 965
- Heavens A., Panter B., Jimenez R., Dunlop J., 2004, *Nature*, **428**, 625
- Hildebrand R. H., 1983, *QJRAS*, **24**, 267
- Hogg D. W., Bovy J., Lang D., 2010, preprint, ([arXiv:1008.4686](https://arxiv.org/abs/1008.4686))
- Hudelot P., et al., 2012, *VizieR Online Data Catalog*, **2317**
- Inoue A. K., Shimizu I., Iwata I., Tanaka M., 2014, *MNRAS*, **442**, 1805
- Jørgensen I., Chiboucas K., 2013, *AJ*, **145**, 77
- Kennicutt R. C., Evans N. J., 2012, *ARA&A*, **50**, 531
- Kewley L. J., Dopita M. A., Leitherer C., Davé R., Yuan T., Allen M., Groves B., Sutherland R., 2013, *ApJ*, **774**, 100
- Kriek M., Conroy C., 2013, *ApJ*, **775**, L16
- Kroupa P., Boily C. M., 2002, *MNRAS*, **336**, 1188
- Law-Smith J., Eisenstein D. J., 2017, *ApJ*, **836**, 87
- Leja J., Johnson B. D., Conroy C., van Dokkum P. G., Byler N., 2017, *ApJ*, **837**, 170
- Lonocce I., Longhetti M., Saracco P., Gargiulo A., Tamburri S., 2014, *MNRAS*, **444**, 2048
- Lotz J. M., Jonsson P., Cox T. J., Croton D., Primack J. R., Somerville R. S., Stewart K., 2011, *ApJ*, **742**, 103
- Madau P., 1995, *ApJ*, **441**, 18
- Maiolino R., et al., 2012, *MNRAS*, **425**, L66
- Maraston C., Pforr J., Renzini A., Daddi E., Dickinson M., Cimatti A., Tonini C., 2010, *MNRAS*, **407**, 830
- McCracken H. J., et al., 2012, *A&A*, **544**, A156
- McLure R. J., et al., 2011, *MNRAS*, **418**, 2074
- McLure R. J., et al., 2013, *MNRAS*, **428**, 1088
- McLure R. J., et al., 2017, preprint, ([arXiv:1709.06102](https://arxiv.org/abs/1709.06102))
- McLure R. J., et al., 2018, *MNRAS*, **477**, 101
- Mobasher B., et al., 2015, *ApJ*, **808**, 101
- Moresco M., et al., 2010, *A&A*, **524**, A67
- Mortlock A., McLure R. J., Bowler R. A. A., McLeod D. J., Marmol-Queraltó E., Parsa S., Dunlop J. S., Bruce V. A., 2017, *MNRAS*, **465**, 672
- Moustakas J., Kennicutt Jr. R. C., Tremonti C. A., 2006, *ApJ*, **642**, 775
- Nelson D., et al., 2017, preprint, ([arXiv:1707.03395](https://arxiv.org/abs/1707.03395))
- Nogueira-Cavalcante J. P., Gonçalves T. S., Menéndez-Delmestre K., Sheth K., 2018, *MNRAS*, **473**, 1346
- Onodera M., et al., 2012, *ApJ*, **755**, 26
- Onodera M., et al., 2015, *ApJ*, **808**, 161
- Pacifici C., Charlot S., Blaizot J., Brinchmann J., 2012, *MNRAS*, **421**, 2002
- Pacifici C., et al., 2016, *ApJ*, **832**, 79
- Panther B., Jimenez R., Heavens A. F., Charlot S., 2007, *MNRAS*, **378**, 1550
- Papovich C., et al., 2012, *ApJ*, **750**, 93
- Peng Y.-j., Lilly S. J., Renzini A., Carollo M., 2012, *ApJ*, **757**, 4
- Peng Y., Maiolino R., Cochrane R., 2015, *Nature*, **521**, 192
- Pentericci L., et al., 2018, preprint, ([arXiv:1803.07373](https://arxiv.org/abs/1803.07373))
- Pforr J., Maraston C., Tonini C., 2012, *MNRAS*, **422**, 3285
- Reddy N. A., Pettini M., Steidel C. C., Shapley A. E., Erb D. K., Law D. R., 2012, *ApJ*, **754**, 25
- Schawinski K., et al., 2014, *MNRAS*, **440**, 889
- Siudek M., et al., 2017, *A&A*, **597**, A107
- Skilling J., 2006, *Bayesian Anal.*, **1**, 833
- Smethurst R., et al., 2017, preprint, ([arXiv:1709.09175](https://arxiv.org/abs/1709.09175))
- Steinhardt C. L., et al., 2014, *ApJ*, **791**, L25
- Straatman C. M. S., et al., 2014, *ApJ*, **783**, L14
- Straatman C. M. S., et al., 2016, *ApJ*, **830**, 51
- Strateva I., et al., 2001, *AJ*, **122**, 1861
- Strazzullo V., et al., 2013, *ApJ*, **772**, 118
- Thomas D., Maraston C., Schawinski K., Sarzi M., Silk J., 2010, *MNRAS*, **404**, 1775
- Thomas R., et al., 2017, *A&A*, **602**, A35
- Tomczak A. R., et al., 2014, *ApJ*, **783**, 85
- Vogelsberger M., et al., 2014, *MNRAS*, **444**, 1518
- Whitaker K. E., et al., 2011, *ApJ*, **735**, 86
- Whitaker K. E., et al., 2013, *ApJ*, **770**, L39
- Wilkinson D. M., Maraston C., Goddard D., Thomas D., Parikh T., 2017, *MNRAS*, **472**, 4297
- Williams R. J., Quadri R. F., Franx M., van Dokkum P., Labbé I., 2009, *ApJ*, **691**, 1879
- Worthey G., Faber S. M., Gonzalez J. J., Burstein D., 1994, *ApJS*, **94**, 687
- Wuyts S., et al., 2011, *ApJ*, **738**, 106
- Younger J. D., et al., 2009, *MNRAS*, **394**, 1685
- van der Wel A., et al., 2014, *ApJ*, **788**, 28

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.